~~FFC~~/Repl.
UL

Psychological Monographs

General and Applied

504
60

Some Variables Affecting Pattern Conception

By

Eva I. Shipstone

Harvard University

Shipstone

Price $1.50

74
17

Edited by Norman L. Munn

Published by the American Psychological Association, Inc.

Psychological Monographs:

General and Applied

Combining the *Applied Psychology Monographs* and the *Archives of Psychology* with the *Psychological Monographs*

NORMAN L. MUNN, Editor

Department of Psychology, Bowdoin College
Brunswick, Maine

Consulting Editors

ANNE ANASTASI
FRANK A. BEACH
ARNOLD M. BINDER
W. J. BROGDEN
ROBERT R. BUSH
JOHN F. DASHIELL
JAMES J. GIBSON
D. O. HEBB
EDNA HEIDBREDER
FRANCIS W. IRWIN
JAMES J. JENKINS
HAROLD E. JONES
DANIEL KATZ
BOYD MCCANDLESS
DONALD W. MACKINNON
QUINN MCNEMAR
LORRIN A. RIGGS
CARL R. ROGERS
RICHARD L. SOLOMON
ROSS STAGNER

Manuscripts and correspondence on editorial matters should be sent to the Editor. *Psychological Monographs* publishes comprehensive experimental investigations and other psychological studies which do not lend themselves to adequate presentation as journal articles. Major space is given to the author's original contribution; introductory and bibliographic materials, as well as statistical tables and graphs, must be kept within reasonable bounds. Tables, graphs, and appendix materials which deal with detail not essential to adequate presentation of the findings may be made available through the American Documentation Institute—for details of this procedure, see the APA *Publication Manual*. Preparation of manuscripts for publication as monographs should follow the procedure given in the APA *Publication Manual*. Publication in *Psychological Monographs* is free of cost to the author, except in cases where early publication is requested or author's alterations are made in galley proofs.

ARTHUR C. HOFFMAN
Managing Editor

HELEN ORR
Promotion Manager

Correspondence on business matters should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St., N.W., Washington 6, D.C. Address changes must arrive by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee third-class forwarding postage.

COPYRIGHT, 1961, BY THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

Vol. 74, No. 17 Whole No. 504, 1960

Psychological Monographs: General and Applied

SOME VARIABLES AFFECTING PATTERN CONCEPTION

EVA I. SHIPSTONE [1]
Harvard University

THIS approach to the study of concept formation is based on the assumption that there is an implicit conception of structure (patterning behavior) in adult human beings and that this can be classified into hierarchically organized categories. In order to give a more definite conception of what this means, we begin with an annotated protocol recorded for a single subject. The procedure described in this sample protocol is slightly different from that used in the main experiments described below; it was chosen, however, in order to unravel for the reader in some detail the process of strategy construction in concept formation. This technique is replaced by a simpler and more concise one in the experiments to be described later.

AN ILLUSTRATIVE STUDY
(Experiment I)

The subject sits at a table; across from him is the experimenter. The subject has just finished reading the following instructions:

> During the last war an army base was receiving important code information, one message at a time. This information had to be sorted for filing as it was received. Now suppose you were the officer on duty with this important job of receiving and filing this information. You will now be presented with the same information, single message on single card, two cards at first, then one card at a time. You are to study the card and then decide how you are to file it. You can put the cards into as many piles or boxes as you wish and you are free to move around a card from one box to another any time you wish. You will not be limited in time since the decisions you make are very important, but work rapidly. Read these instructions again until you understand them thoroughly and read them again in between if you should forget them.
>
> Any questions?

Now, one further point. As you work with these cards, think aloud as if you were talking to yourself. Talk freely about what you are doing throughout your performance.

Any questions?

"Here," says the experimenter, "are the first two cards." The subject is handed two cards, 5″ × 8″, on which are typewritten: ZIR PAG ZIR PAG NEN; and ZIR PAG NEN PAG NEN PAG NEN PAG NEN PAG NEN PAG NEN PAG ZIR NEN. Only three nonsense words, NEN, PAG, ZIR, arranged in various different orders, are typed on any of the cards; it is convenient to abbreviate them, N, P, and Z. The subject looks at the two cards, laughs self-consciously, and places them separately on the table in front of him.

He says, "It seems as though the main difference in the way things repeat is first according to the symbols that appear and then in the order in which they appear."

There is a silent pause. The experimenter says, "Now you will have to tell me when you are ready for the next card."

"Next card, please."

The next card contains the message ZPZN. The subject places it on the table by itself. All three cards are separately placed, so that the experimenter asks, "Why did you place that card there?"

"Oh, well," the subject says, "there seems to be a difference in the lengths of the messages. And I don't have enough information to really decide on categories. I am just setting them aside until I gather a little more information."

The experimenter, who has seen some people base their entire classification on the length of the messages, makes a special note that the first basis of classification mentioned by the subject is length. Will he stay with this aspect, or will he look further?

The next card is ZPZZZZZZZZZZZZPN. The subject says, "Now this card is more similar in length to one of the other ones I received, and it contains a large number of ZIRs as the other

[1] The author, now teaching at Isabella Thoburn College, Lucknow, India, wishes to acknowledge her indebtedness to George A. Miller for directing this research. Without his invaluable help this thesis could not have been written. Gratitude is also expressed to E. B. Newman, Chairman of the department, for his suggestions.

one, so I will put these two together, possibly in the same category." He places together ZPZZZZZZZZZZZZPN and ZPZZPN; since he does not combine ZPZZZZZZZZZZZZPN and ZPNPNPNPNPNPNPZN, he seems to have rejected length as a basis of classification. For the moment, apparently, his attention is caught by another aspect, the repeating unit ZIR.

The next card is ZPNPZN. The subject reads it silently. There is a long pause. "Whatever you are thinking," says the experimenter, "please say it aloud. I know that is not easy."

"Well," says the subject, "I just noticed that all of the five cards start with the same syllable and the syllables appear to be ordered. Similarly, the second syllables in all the cards are the same. A difference arises in the third syllable, so there is a possibility of arranging these cards according to the syllables. Alphabetically, in other words. But for the time being I will put this card in with the one which has similar length, that is, with the same number of syllables, like six or so." He proceeds to file them by length, as follows: ZPNPZN and ZPZZPN are put together, ZPNPNPNPNPNPZN and ZPZZZZZZZZZZZZPN are put together, and ZPZN is left by itself. Although he is beginning to pay attention to the way the messages begin, and to consider the possibility of an alphabetical ordering, he rejects it in favor of a classification on the basis of length. He has destroyed his first tentative grouping on the basis of a repeating unit and has returned to the earlier hypothesis.

The next card is ZPPN. "Again," says the subject, "all the cards start with the same two syllables. But I am going to set this card aside for a separate category for no other reason than that I want to gather more cards before I make a classification." There is a pause. "Am I to assume that this is coded information? I mean, information like language information?"

The experimenter replies, "I am sorry I cannot answer that question for you. These are messages you are receiving and you are trying to file them."

"Messages are coming in and I am trying to organize them. Hmm . . . temporarily, I am going to rearrange the cards and put them in alphabetical order according to the syllables, starting from the beginning." The subject seems to hope that something not presently obvious will appear as he tries different spatial arrangements of the cards. As he works at alphabetizing them, he says, "I've just noticed something that had not occurred to me before, that there are only four symbols." In fact, of course, there are only three. "Somehow," he continues, "in the beginning I assumed there were going to be more. So, now I am arranging them first of all according to syllables, that is, the syllable order and the first three syllables, and then according to the lengths of the messages." Since the first two syllables are identical on all the cards that he has seen, any attempt to alphabetize must take into consideration at least the first three syllables of the message.

The experimenter notes that the subject is beginning to combine his rules in an effort to find some new basis for classification. Like most subjects, this one began with relatively simple rules—counting and alphabetizing—ready-made operations he has learned in other contexts; he transfers them bodily to the present task. For some reason, however, the results do not satisfy him and he is now beginning to search for something else. As he collects more cards, he wants to come to terms with them in a way that respects their peculiarities and particularities, not by some ready-made rule that might fit almost anything.

As shown in the chart below, the next card is ZPNPNPNPNPZN. The subject studies this card in silence for almost 60 seconds, looking back and forth between it and other cards spread out on the table before him. Then he makes a new observation, "I notice now that when I look at these cards in categories that there are units within each card that seem to repeat between cards. Like, for instance, I have two cards here both of which start with ZIR PAG NEN and end in ZIR NEN, and in between they have the unit PAG NEN, but in slightly different numbers. I mean, one card will have more of these units than another card. It seems like relationships are starting to emerge."

The more complex rule that the subject has just discovered is one that the experimenter has watched many subjects come to; the experimenter calls such formulations "Initial-Middle-Final Rule." Although the rule is here applied to only two cards to form a single category, it contains the germ of the filing system that this subject will use for all the remaining cards. Note that it is a rule designed to fit the present situation and to use the particular information provided by this set of messages; it is not a rule that would apply equally well to all imaginable sets of messages.

The order in which the messages were presented to the subject was:

1. Z P Z Z P N
2. Z P N P N P N P N P N P N P Z N
3. Z P Z N
4. Z P Z Z Z Z Z Z Z Z Z Z Z Z P N
5. Z P N P Z N
6. Z P P N
7. Z P N P N P N P N P Z N
8. Z P Z Z Z Z Z P N
9. Z P Z Z Z Z Z Z Z Z Z Z Z Z Z P N
10. Z P P P P N Z N N N N N P
11. Z P P P N Z P
12. Z P N Z N P
13. Z P N P N P N P Z N
14. Z P P N Z P
15. Z P Z Z Z Z Z Z Z Z Z N P
16. Z N Z N N N N P
17. Z P N P N P N P N P N P N P N P Z N
18. Z P Z Z Z P N

19. ZPZZZZZPN
20. ZPNPNPNPNPNPNPNPZN
21. ZPPPPPPPPPPPNZNNNP
22. ZPZZZZZZZPN
23. ZPZZZZZZZZZPN
24. ZPNZNNP
25. ZPNPNPNPNPNPZN
26. ZPZN
27. ZPNPNPNPNPNPNPNPNPZN
28. ZNZP
29. ZNZNP
30. ZPPPPPNZP

At this stage the subject has received seven cards and has arranged them into three sets as follows, which we shall label as A, B, C, for convenience:

A. ZPZZPN
ZPZZZZZZZZZZZZZPN
B. ZPNPNPNPNPZN
ZPNPNPNPNPNPNPZN
C. ZPPN
ZPZN

The seventh card, ZPNPZN, is left isolated for the time being.

The next card is ZPZZZZZZPN. "Again I see the same thing within cards," the subject says, apparently with some satisfaction. "I see units repeating in the cards within each class. For example, there is an end unit, two end units, in fact, and the intermediate units are the same, just varying in number. In other words, I count the number of ZIRs in the middle unit in both cards. It may be five in one case and six in another, and this holds for all my cards, I think. In other words, in all cases the ends and the beginnings are the same. It is the middle that varies."

The next card is ZPZZZZZZZZZZZZZPN. The subject says, "This classification seems to be working out very nicely, so I will continue."

The next card is ZPPPPNZNNNNNP. There is a long silent pause; clearly, something new has appeared. The experimenter reminds the subject to keep talking.

"All right," he says, "now we have another card. I am trying to see what category this will fit in. I am looking at the first three syllables and again I see the same pattern emerging, I think. I notice now, as a side observation, we seem to have only three basic categories. And each of these categories is divided into subclasses which consist of cards having the same beginnings and endings and the same intermediate units, varying in number." The subject places this card, ZPPPPNZNNNNNP, together with ZPPN and ZPZN.

The next card is ZPPPNZP. "Again, the same pattern. I mean to collect more cards and see if it continues. And as soon as I get a few more cards to form my third category I am going to see if there are any more relationships observable between the major categories themselves."

The next card is ZPNZNP. "Oh," says the subject, "now I notice that I made a mistake. I see that whereas I thought there were three categories I now see there are four. The end units of two of my cards in one of my classes is different." He indicates Set C consisting of ZPPN, ZPZN, ZPPPNZP, and ZPPPPNZNNNNNP. "In fact, there are considerably more categories than I thought. My first category (Set A) is all right. The beginnings and ends are the same, and the intermediates are the same. Similarly, my second category (Set B) is all right. Just this one I was trying to make my third category seems to be wrong. It does not seem to follow the same pattern the other two do. I just assumed it did, but I see that it doesn't. I could attempt to follow the same classification, that is, instead of having three categories, I could have five or six. In other words, I will continue to see if a pattern emerges. It looks like I will continue with this new category. It appears, though, that I may have to discard my old standards, or criteria."

The subject has been given no hint as to how many categories there should be; concern with the number is his own attempt to define the terms of the task. The desire to have as few categories as the evidence will allow seems to be an important consideration governing many subjects' satisfaction with their own efforts. Moreover, like most subjects, he wants to use the same rules for every category; it feels inconsistent, or at least untidy, to have a different description for each different set of messages.

Note here, incidentally, that it is his third category with a double variation in the middle that is causing his troubles.

At this stage the subject has received 12 cards and has arranged them as follows (though not necessarily in the same order within classes):

A. ZPZZPN
ZPZZZZZZPN
ZPZZZZZZZZZZZZZPN
ZPZZZZZZZZZZZZZPN
B. ZPNPNPNPNPZN
ZPNPNPNPNPNPNPZN
C. ZPPN
ZPPPNZP
D. ZPNZNP
ZPPPPNZNNNNNP

Two cards, ZPZN and ZPNPZN, are left as odd cards, not yet assigned to any definite category.

The next card is ZPNPNPNPZN. The subject says, "I just noticed something I hadn't noticed before: that all the cards begin with ZIR." Has he rediscovered something he remarked on first after the fifth card? Or is he noting that his earlier generalization is still valid? He says, "There no longer seems to be the tendency to have the same end units." Perhaps he means that the beginnings of the messages are still predictable, even though the endings are beginning to seem more variable. Unfortunately, it is not always

possible to understand what a subject means; to interrupt him by questions would interfere too much with the processes the experimenter is trying to observe.

The next card is ZPPNZP. "I noticed that the card I just received is like the first card in one of my categories, the simplest card, only this has additional symbols at the end." He places ZPPNZP together with ZPPN and ZPPPNZP. "This could not occur with any of my other cards," he says. "I will continue." He means that the endings of two of his cards in a category could be different.

The next card is ZPZZZZZZZZZNP. "Now this card fits into one of my earlier categories. I will put it there."

The next card is ZNZNNNNP. There is a long pause. Then, "Now I am lost."

"Why are you lost?"

"I don't see relationships occurring between these last five or six or seven or eight cards. I cannot see where I can fit them sensibly into the categories. They don't seem to be fitting into the same nice pattern that the earlier cards did." He leaves this card unassigned to any of the groups.

The next card is ZPNPNPNPNPNPNPNPZN. "This one fits into one of my earlier categories. What I would like to see . . . I would like to get more information in order to decide whether or not I have to abandon my older classification. Maybe these cards that I have will fit better if a card could be used to establish a new category using the same criteria as before. It is not apparent, however, whether this will be useful in sorting these materials. We have sort of new materials here. The two categories that I have seem to be filling in very nicely; they seem to be complete series, whereas these other cards don't fit anywhere, yet."

The cards that are giving our subject so much trouble all belong together; the confusion is arising because there are two variables in this set as opposed to only one variable in each of Sets A and B.

While the subject was talking he reached for another card, ZPZZZPN, which he placed in one of his earlier categories.

The next card is ZPZZZZZPN. He greets it like an old friend and disposes of it quickly. "Now this one, too, fits into my older category, so I will place it there."

The next card is ZPNPNPNPNPNPNPNPZN, another old friend. "And, similarly, this one fits in all right."

The next card is ZPPPPPPPPPPNZNNNP. A stranger, alas. "Now, I have this card. . . . I see another pattern emerging, I think. . . . There are . . . now, let me think . . . yes, it's just that I didn't discover what the intermediate unit was in some of these, assuming it to be two syllables, like the more obvious one in the other set," he says, comparing this card with the one just preceding it. "I see now that the end syllable consists of just one—I mean, the ending consists of just one syllable, and the intermediate ones are more than one syllable. I can form another classification, using the beginning three symbols and ending in one symbol, at least for some of my cards. I can fit them into my classification. In other words," he says, revealing once more his desire to find a consistent basis for the classification, "I have formed a third category which is similar to the other two larger categories I have (meaning Sets A and B). And I still have four cards left over." He refers to ZPPN, ZPZN, ZPNPZN, and ZNZNNNNP as the four that are left over; then, returns to his new, or third, category in search of a simple rule: "Now, let's see if we can see the same sort of relationship here. Ah, it seems to be emerging. There is an end unit which is constant and there seems to be a middle unit which seems to be constant. So we can have two ends and a middle; then, in between ends and the middle, there are other units which are the same. The same unit is added, only just in different numbers. So we have the same sort of process as before, except a little more complex. Now the question is, how to classify these cards. We will make a category out of all the cards that have the same two end units and the middle unit and I will classify them in order of their relationships depending upon the number of units in the beginning, and then the number between the end and the middle. If there is one unit, it will be the first card; if there are three units, it will be the second card, assuming that there isn't one with two units in between. And if I have two cards with the same number of units repeating in between the beginning and the middle units, then I will use the second side—I mean, the number of syllables that is between the middle and the end is my second way of classification. I will rearrange these now."

Notice how the subject distinguishes between variant and invariant units in the statement of his rules. The approach might now be characterized as Initial-x-Middle-y-Final, where x and y represent the variable parts. And he also has the rule well enough in mind to use it for ordering the cards within each category.

"When I come to rearrange these," the subject says, "I see this isn't the case."

"What isn't the case?"

"That the two ends and a middle unit are constant with all the cards, and there is an extra syllable. So, I am going to wait."

At this stage the subject has received 21 cards and has rearranged them as follows (though not necessarily in the same order within classes):

A. ZPZZPN
ZPZZZPN
ZPZZZZZPN
ZPZZZZZZPN
ZPZZZZZZZZZPN
ZPZZZZZZZZZZZZPN
ZPZZZZZZZZZZZZZZPN

B. ZPNPNPNPZN
ZPNPNPNPNPZN
ZPNPNPNPNPNPZN
ZPNPNPNPNPNPNPZN
ZPNPNPNPNPNPNPNPZN

C. ZPNZNP
ZPPNZP
ZPPNNPP

ZPPPPNZNNNNNP
ZPPPPPPPPPPNZNNNP

Four cards, ZPPN, ZPZN, ZPNPZN, and ZNZNNNNP, remain unclassified. Although he is making progress, he does not feel completely confident about the third category. [At this point the reader should try, if he has not already, to fit the four odd cards into the existing categories; he will discover that this subject, like the majority of subjects, finds it difficult to invent zero when he comes to count the number of times a unit recurs.]

The next card is ZPZZZZZZZPN. "This one fits into one of my older classifications, and in my first big category."

The next card is ZPZZZZZZZZZPN. "And this one, too. I will put it there."

The next card is ZPNZNNP. "Now I notice these . . . here . . . no, I'll see if it fits into one of my categories and if not, I will put it with these other cards and see if I can establish a new category with them. Now, this last card seems to be similar to two of my cards that I had set aside." He puts ZPNZNNP together with ZPPPPNZNNNNNP and ZPPPPPPPPPPNZNNNP. "They have the same beginning and the same end and pretty much the same middle unit, but not quite, however."

The next card is ZPNPNPNPNPNPZN. "This fits into one of my older categories."

The next card is ZPZN: it is the same as the third card he had received earlier [pause]. "This one I will place with that and that," he indicates ZPNZNP and ZPNZNNP. "There's . . . [a pause] it just occurred to me that, although there appears to be a relationship, it is probably only by virtue of the fact that there are only four symbols we are using." The mistake persists; he still does not realize that, in fact, only three symbols are used. "However," he continues, "there does appear that there is an order. I mean, if I were to divide it like I have two end units and middle units, it appears that there is another syllable appearing in between the beginning and the middle and it is always the same unit. It just varies in number, much like the other classification. So now I will order those. As a matter of fact, I see that three of my cards fit very nicely, actually, there are four that fit in very nicely. Five . . . oh, no, this is a little different. The end unit is different, so I have now established the fourth category. And I still have two cards left over that I am not quite sure what to do with. Let me check now and see if they don't fit somewhere . . . no, they don't seem to fit."

The next card is ZPNPNPNPNPNPNPNPNPZN. "This fits in very nicely with one of my earlier categories." He disposes of it quickly.

The next card is ZNZP. "This card does not seem to fit anywhere."

The next card is ZNZNP. He pairs it with ZNZNNNNP, saying, "Now I have two odd cards that seem to fit to form another category, using the same idea of picking out syllables interspaced in between, just repeating. There is just one thing I want to say pertaining to one unit. . . . So, I now have five categories and two extra cards which, temporarily, I will call separate categories."

The next card is ZPPPPPNZP. The experimenter says, "This is your last card."

The subject says, "This one, I think, fits into one of my earlier categories. Now I want to look over these couple of categories and see how I stand on those. I want to make sure that I have really followed the rules I think I established."

"You can take all the time you want to," says the experimenter, "but don't forget to talk."

At this stage the subject has received 30 cards and has arranged them as follows (though not necessarily in this order within classes):

A. ZPPN
ZPZZPN
ZPZZZPN
ZPZZZZZPN
ZPZZZZZZPN
ZPZZZZZZZZZPN
ZPZZZZZZZZZZPN
ZPZZZZZZZZZZZZPN
ZPZZZZZZZZZZZZZZPN

B. ZPZN
ZPNPZN
ZPNPNPNPZN
ZPNPNPNPNPZN
ZPNPNPNPNPNPZN
ZPNPNPNPNPNPNPZN
ZPNPNPNPNPNPNPNPZN
ZPNPNPNPNPNPNPNPNPZN

C. ZPNZNP
ZPNZNNP
ZPPPPNZNNNNNP
ZPPPPPPPPPPNZNNNP

D. ZPPNZP
ZPPPNZP
ZPPPPPZP

E. ZNZNP
ZNZNNNNP

Two cards, ZNZP and ZPZN, remain unclassified.

"Okay," the subject says. "Well, I'll start out with my biggest and main categories because the relationships are clear, so I'll make sure that they

follow all the way through. Now, they all have the same beginning and end syllables. There is no question about that. And, similarly, the second and next-to-last symbol in all—the beginning and end units are similar in all cards. In my first category the intermediate symbols are always ZIRs and they occur in varying numbers. Let's see, one, two, three, skip some; I can order these according to the number of ZIRs in between. I don't suppose it matters as long as I am classifying rather than trying to interpret the information. I can't seem to decide between these: they seem to be homogeneous."

He shifts to Set B. "I'll take the next one. Again, we have two syllables that are the same—I mean, they are the same from card to card, beginnings and ends. And we have in between them, rather than one syllable repeating, a pair of syllables repeating . . . and always ending in the same one. So I order these according to the number of insertions."

He shifts to the Set C. "Now I am going to look at one of the more complicated categories. Here it seems there's a beginning unit that consists of two syllables on each of the two cards and the end unit which consists of only one syllable" [short pause]. "In addition, there is a middle unit that consists of NEN ZIR, or whatever it is. And in between there are varying numbers of syllables between the beginning units and the middle unit. In this category there are varying numbers of syllables PAG, so I am going to arrange them in this order. This will be my primary ordering process. And, should there be duplication in the number of PAGs, then I would check the number of symbols on the other side of the middle part, and order them according to their number" [pause]. "I see that this does occur . . . so, they are ordered first with respect to the symbols between the beginning and the middle, and then with respect to the symbols between the middle and the end. Should there be any ambiguity in the first order . . . it appears like a very nice. . . ." He combines his next three categories into one on the basis of the rule just stated; the result is a third category:

C. ZNZNP
ZNZNNNNP
ZPNZNP
ZPNZNNP
ZPPNZP
ZPPPNZP
ZPPPPNZNNNNNP
ZPPPPPNZP
ZPPPPPPPPPPNZNNNP

"Now," he continues, "I have the cards in three main categories, and I'll see if they will hold up. Again, they begin with, and end with, the same unit; and the same pattern of middle syllables repeats."

He shifts to the remaining odd cards, ZNZP and ZPZN. "Now let's see if my last category. . . . I'm not sure what to do with that—make sure if it doesn't fit into one of the old ones . . . it doesn't as such . . ." [long pause]. I'm glancing over all the categories now to make sure that this last category doesn't fit into a subset of those" [pauses while looking]. "I think. . . ." He picks up ZPZN. "I now notice that one of these cards does fit in to the, ah, my earlier category. . . . It's curious, I didn't notice it before, but it seems to be identical with the first one in one of my categories. Well, I'll place it there."

There is a long pause while he considers the remaining odd card, ZNZP. "It seems a shame to have a one-card category. I'll look over once more to see if we can find a place to put it . . . this odd card."

After a little more search he gives up. "I will leave it as it is," he says, "I am happy with my classification."

The above technique was abandoned in all the studies to be described below in favor of a simpler and quicker one: all the cards were simultaneously presented to the subject and no running protocol was recorded as the subject sorted the strings Comments the subject wished to make on his method of sorting were recorded only at the end of the experimental session.

The Problem

This account of a subject's responses to a problem solving situation has been reconstructed from a tape recording. The subject is confronted with the problem of processing information which he received a little at a time. We can think of the subject as starting the task in total darkness, for he has almost no idea what kind of information is encoded in the messages. Whatever he does with the messages is entirely a matter of his own decisions.

After watching a dozen or more subjects perform in a similar manner, one is impressed with several things. The first is the similarity between this laboratory episode and ordinary behavior shown in dealing with our daily environment. All day long stimuli impinge upon an organism in a haphazard array; unless he can rapidly sort and file information in the right compartments his life will quickly revert to chaos—to the "buzzing, blooming confusion" of the newborn child. Man orders his environment according to rules that he develops for himself or which he learns

from others; by this recoding technique he can economically handle the multiplicity and the variability of his environment for present and future use.

A second impression one gathers is that men have learned to behave in this orderly manner without any specific instructions and without any knowledge of results provided by a teacher or an experimenter. Man seems to respect structure in the world wherever he finds it, regardless of consequences. It does not seem necessary here to invoke any hedonic contingencies—unless perhaps orderliness carries its own rewards: a satisfaction with the goodness of fit. The subject is uneasy when a string of symbols does not fit one of the categories he has evolved. What one means by the fit here is more subtle, and the lack of fit less obvious, than in a mechanical or a picture puzzle; without any outside guidance or control the subject moves from a search for isolated perceptual relations into a search for the conceptual pattern underlying the stimuli that he receives.

The behavior of our subject emboldens us to argue for an implicit conception of structure in man. The argument is, of course, not new to psychology. It was first brought into lively focus by the Gestalt psychologists as part of their protest against the Wundtian tradition of elementism and against the behavioristic denunciation of everything mentalistic. The Gestalt psychologists did not share with Behaviorism its contempt for mentalistic concepts, nor its distaste for introspection. On the contrary, they were against any oversimplified mechanistic view of man; they urged the importance of value in a world of facts (Köhler, 1938). Meaningful relations (Köhler, 1929; Wertheimer, 1945), organization (Köhler, 1929), fitness of parts into an organic whole (Koffka, 1935), requiredness (Köhler, 1938)—these were some of the values they wished to restore to a world hacked up into elements.

The emphasis on relations is not a novel thought, either. Even Thomas Brown (cited in Woodworth, 1948) early in the nineteenth century, and von Ehrenfels (1937) later in that century, had stressed the importance of relations. But the notable contribution of the Gestalt school was to demonstrate the crucial role of structure, or pattern, both in the stimulus configuration and in the responses of the observer. According to them the perceptual process is essentially one of organization. The brain is a dynamic neural center where perceptual objects combine or separate on the basis of similarity, contiguity, symmetry, common fate. Köhler's conclusion (1947) that an "organism responds to the pattern of stimuli to which it is exposed" (p. 103) seems to fit the illustrative account of our subject's responses to a decision making task. We cannot help but notice how often the subject is concerned with the "fit" of the messages into some category he had established: he is continually seeking "relationships," he talks about "pattern," he is trying to fit the parts into a total scheme he has set up to solve his problem. In fact, Wertheimer (1945) might have been describing our own subject instead of the children who solved geometry problems, when he wrote: "In all these cases we have structural changes, tendencies toward the better structure, toward fitting, with the disappearance of disturbances" (p. 57).

Since the notion of structure, or pattern, as proposed by the Gestalt psychologists, seems essential for the description of our subject's behavior, it becomes the starting point of our inquiry. The notion of "pattern," however, as we use it here, is something of a special case of the Gestalt notion. We are interested in what Miller and Chomsky (1958) have called "pattern conception." For the purpose of the present experiments we define pattern as the invariance of a sequence of symbols in the context of lawful change, and we simplify its meaning by defining pattern in terms of internalized rules, or relations, a subject employs to organize novel information. Rules usually imply a notion of abstraction and generalization; they go beyond the information given. They cannot be said to be covered either by the transposition experiments such as those of Köhler (cited in Osgood, 1956, p. 495), nor by the studies of stimulus generalization conducted by the

Behaviorists (Hovland, 1937a, 1937b; Razran, 1940, 1949). In both cases the results could be explained on the basis of sensory relations alone. By the use of rules the subject searches for pattern in order to recode the sensory input into manipulable, meaningful units. Perhaps a limitation on the information capacity of the human channel makes such behavior (recoding) necessary. In any case, it is our wish to find out something about these rules or patterns by which the subject handles information. We would like, therefore, to make a systematic study of some of the experimental variables that affect the nature of the rules that subjects employ in order to search for pattern in the stimulus input.

One could begin a study of rules by asking the usual questions of What? How? and Why? Hull (1920) is one among those researchers who tackled the question What? "What does a child learn," he asked, "when he calls a certain object dog?" This question is part of a broader question: What does any individual learn when he is said to have learned a concept? To find an answer Hull required adult human subjects to learn some Chinese sounds paired with certain characters in the Chinese alphabet. Hull's conclusion was that *what* an individual learns is a rule about the common characteristics of a class of objects. He placed the emphasis on the common characteristics of the stimuli; the rule, Hull equated with discrimination learning. It is, therefore, not clear to the reader whether Hull's subjects learn a rule or whether they merely learn a name for the rule.

A second answer to the question *What?* is contained in the word relations. According to a second view what the subject really learns is some kind of a rule of relations by which he groups instances as members of a class. A search for common relations among instances might take the subject beyond the common perceptual elements to something more. For the Gestalt thinkers this something more is the "organic whole" which is more than the sum of its parts. Goldstein and Scheerer (1941) describe it as "abstraction." According to them the process of abstration is different from that of aggregation of common elements under one label. Smoke (1932) discusses this same idea as symbolic reaction. Our own approach to rules is in terms of relations and we view rules of relations as pattern conception.

The question—how rules are learned or made—has been ingeniously handled by Bruner, Goodnow, and Austin (1956), and more recently by David Reiss (1958). Bruner et al. systematically investigated the methods by which subjects categorize given events. The categories into which events are to be placed are predetermined by the experimenter's rules. It is the task of the subject to discover the experimenter's rules for placing the given events into their proper boxes. In order to categorize successfully the subject uses "strategies," or plans, by which he searches for information which would be useful to him in constructing his hypotheses about how the events are distributed into the experimenter's boxes. This stepwise process—of strategy construction, searching for information, hypothesis building, testing hypotheses—the authors describe as the process of concept attainment. Of these steps, the one of strategy construction is the most critical. Bruner et al. (1956) describe strategy as "regularities in decision-making," and define it as "a pattern of decisions in the acquisition, retention, and utilization of information that serves certain objectives (p. 54). A "host of factors" influence the construction of strategies in concept attainment, according to these researchers.

In his discussion of the construction of strategies by subjects, Reiss takes Bruner's "host of factors" for granted, and he proceeds to study how a subject constructs his "model" of what the rules that govern the stimuli might be, but Reiss studies this without constraining the subject strictly by the experimenter's rules. He varies the quantity of evidence given to the subject initially, and then observes how this amount of information affects the subject's behavior in model attainment. Reiss finds that the subject's model is in fact seriously

influenced by the amount of information initially presented to the subject.

In all the studies discussed above, except the last one, the experiment concludes when the subject reaches the solution as the experimeter stipulates it. Hull's subjects continue to learn the list until the criterion is reached. The criterion, of course, means the one that the experimenter has set up. Hanfmann and Kasanin (1937) turn over a block or two to suggest to their subject the "correct" grouping which he must discover. Bruner, Goodnow, and Austin confirm and infirm hypotheses. In none of these studies is the subject ultimately free to set up a criterion independently of the experimenter. Obviously then, the question arises whether the subject will proceed to work by rules even when he is free from the experimenter's criterional constraints.

Thorndike (1931) asked subjects to draw lines equal to a specified length without giving them any information about the accuracy of their results. Yet one notes in his results that the performance of subjects began to stabilize around some standard implicit in the subjects' behavior. In other words, the subject seems to be settling on some rule which he bases upon his own performance. It is immaterial whether he does this consciously or unconsciously.

One notices a similar reliance upon a subjective standard in Reiss' subjects as they construct models. The "random" group, whose hypotheses were confirmed or infirmed according to a prearranged random schedule, did not perform randomly at all. Instead, subjects constructed models quite early and began to test these on the basis of their own hypotheses, disregarding the information they received from the experimenter.

Systematic responses, which may be abruptly shifted to new modes of systematic behavior have been observed by Krechevsky (1932), Witkin (1940), and Spence (1939) in animal subjects. By means of a four-point discrimination apparatus, Krechevsky presented an "insoluble" problem to his rats. These subjects did not run around the apparatus in random fashion, however; they developed systematic modes of responses. Later, Witkin, substantiated these results on animals run in nonspecific problem situations, but his explanation was different. Spence observed the same phenomenon in chimpanzees when he presented them with an insoluble problem in a multiple-choice apparatus.

We have mentioned these studies for three reasons: first, because all of them employ a notion of a criterion of success which is externally determined; second, that this notion of an externally imposed criterion introduces some contingencies of reinforcement for the subject; and third, that when the contingencies of reinforcement occur haphazardly, the subject disregards these and begins to develop some stable modes of responses on the basis of his own performance.

All of these observations influence the design of our research; we deliberately take our point of departure from the first two and confine our interest to the third. We stated earlier that our interest is in rules. What kinds of rules does a subject make? How are his rules different from the experimenter's rules? Do these rules depend upon the nature of the stimuli, upon a knowledge of results, or upon the subject himself, or on a combination of these factors?

These and a host of other questions clamor for recognition and investigation. Twentieth century psychology has not been indifferent to this need. We have already seen how some of the earlier studies, characterized as dealing with concept formation, were investigating rules of induction. Within the same framework, but one step further, Heidbreder (1946) introduced the idea of hierarchy in rules. The notion of hierarchy of rules implies the notion of complexity of strategies that subjects construct. In later studies the researchers were interested in the effect of rules upon learning and recall. For example, Miller and Selfridge (1950) tested their subjects on memory of passages which were graded ac-

cording to their statistical approximations to English from zero to high order. The better recall of meaningful materials, the authors explained on the basis of the subject's familiarity with the rules that governed the sequential properties of the meaningful materials, not the fact of meaning itself.

Miller deals with the facilitatiing effect of rules of formation upon learning and recall in a later study (1958) in which he compares the free recall of redundant and random strings. Again he finds that even when the subject is unaware of the rules of formation that govern the structure of redundant strings he still does better in memorizing the redundant strings. Prior to this study of Miller's, some experiments had been done by Aborn and Rubenstein (1952; Rubenstein & Aborn, 1954) to study the effect of rules of contextual constraints upon learning sequences of nonsense syllables. Test passages were developed by progressively limiting the context in which the nonsense syllables could appear. Thus the rules ranged from purely random to highly redundant. These researchers found that when subjects were made familiar with the rules of formation, redundant sequences were somewhat easier to learn, but the amount of information contained in the recalled materials was not constant from passage to passage.

So far the rules we have referred to have been simple rules; they pertain to two kinds of stimuli—unit stimuli and sequential stimuli—both of which have been generated with some kinds of rules of formation. Language is a good example of the second type of stimuli. Miller's rules, as well as those of Aborn and Rubenstein, and of Reiss, deal with rules of this type, that is, with rules of formation that govern sequential dependencies.

Rules that govern sequential properties of a language also have important effects upon its statistical properties (Miller, 1951). Shannon (1949) analyzes the sequential properties of language by probability rules, a technique that provides access to information measurements as well as to organization. His rules deal with the selection of a message from a determinable number of alternative messages that could occur, and with the probability of its occurrence. His rules deal with the amount of information contained in a message. Since Shannon's concern is primarily with the communication of a message and not with its use, value, storage, meaning, history, or anything else, his rules are not of direct interest to us in the present context. His terms, however, and his techniques for measuring the amount of information transmitted through a noisy channel provide the psychologist with some useful tools of analysis which information theorists (Attneave, 1959; Hovland, 1952; Luce, 1954; Miller, 1953; Quastler, 1955) have applied for measuring organization. We, too, shall draw upon these aids.

In contrast to Shannon, the communication engineer, we can place Chomsky (1957), who looks at language as a grammarian. His rules are more general and more formal than those of Aborn and Rubenstein, though not more difficult; he refers to his rules as "finite state grammars." These simpler grammars enable us to avoid the complications of natural languages and to work with simpler structures. But although his notion of rules is formally simple, it raises a few complex questions regarding relation between rules and performance. Miller (1958) points to this problem and he would "like to know what changes, if any, in a set of rules leave performance invariant," (p. 485). And so would we.

Simple rules of formation of sequential stimuli do not always produce simple effects upon responses. Some rules facilitate performance (Anderson & Leonard, 1958; Fitts, Weinstein, Rappaport, Anderson, & Leonard, 1956; Miller, 1951) and some rules hinder it (Anderson & Leonard, 1958). It seems necessary, therefore, to search for parameters that affect the construction and use of rules of formation. It is the purpose of this study to explore some specific experimental variables implicit in these rules that influence the patterning responses of subjects.

The role of specific stimulus features in *pattern perception* has been previously studied by the Gestalt group, by Attneave (1954, 1955), by Fitts and his colleagues (1956), and by others (French, 1953, 1954b, 1954c). Gestalt psychologists analyzed the peripheral properties in visual organization as proximity, similarity, continuity, and closed figure, but the process of organization itself could not, according to them, be understood without the field concept. They postulate an isomorphism between visual organization and the brain processes. According to the principle of isomorphism the environmental field tends to become "psychologically organized" in terms of regularity, symmetry, and simplicity. Attneave translated and quantified the Gestalt factors of organization in terms of information measure of redundancy in the visual input. He measured memory for simple dot patterns on recall. He found that good form, redundancy, and noise were some of the experimentally accessible variables that affect pattern perception. Fitts et al. at Ohio State University have done special work with redundant forms in visual pattern recognition. They developed metric figures of random and constrained type and they measured the amount of redundancy contained in these figures. They studied the correlation of a measured amount of redundancy with recognition of visual forms. Some additional variables, such as language habits and context (Miller, 1951), also have been found to affect the probability of response. Carmichael, Hogan, and Walter (1932), for instance, demonstrated the influence of instruction upon response. Although we shall give a different interpretation of their results later, we shall not ignore the importance of instruction as a possible source of variation in performance.

All the studies mentioned above suggest that a quantitative analysis of the experimental variables in pattern conception is possible. These studies also help to separate some parameters of pattern conception, particularly those concerned with the number of different alternatives used and with structural similarity among the sequences. The number of different alternatives has proved to be an interesting parameter in studies in perception (Attneave, 1954) and in memory (Brogden & Schmidt, 1954a, 1954b). Whether it is an important variable in pattern conception as well is the question we raise. Does a larger vocabulary help with the better ordering of the sequences of symbols or not? If not, why not? One might expect the facilitatory effect of the larger number of different alternatives to reach an asymptotic value at some point.

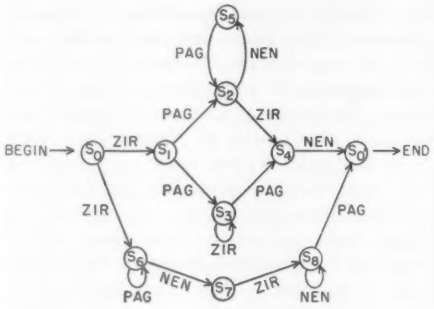
Structural similarity is another parameter that must be explored. Its importance in visual pattern recognition (French, 1954a, 1954b), and in learning and recall (Aborn & Rubenstein, 1952; Miller, 1958; Rubenstein & Aborn, 1954) has been demonstrated. Does it also affect pattern conception? The notion of structural similarity as used in this study should not be confused with Shannon's notion of redundancy. The languages we are working with are, in a sense, "degenerate" because in the limit their redundancy is 100%. We define structural similarity as similarity in rules of formation and specify it in terms of the number of basic strings as well as the number and placement of recursive loops within each string.

Materials and Procedure

Miller and Chomsky (1958) have pointed up some of the difficulties of method and material in laboratory studies of concept formation. In the earlier concept formation studies the subject was constrained by the nature of the task to find the rules as the experimenter defined them. Furthermore, the use of geometrical patterns differing in size, shape, and color effected a certain stereotypy in material and usually limited them to a finite number of possible alternatives from which the positive instances were to be selected. Miller and Chomsky proposed the use of a new kind of material in the study of pattern conception, viewing it as a different kind of task: instead of spatial patterns, they proposed to study sequential patterns. By using finite state systems for materials one could

study pattern conception in a situation "modelled after the task of a grammarian."

A finite state language is an artificial language consisting of the set of well-formed strings (sentences) that can be generated by a set of simple rules (see Chomsky & Miller, 1958). This set of rules is a kind of algebra, or grammar, for identifying the grammatical sentences in a particular language. A finite number of symbols makes up the vocabulary of a language. A finite state grammar consists of a finite number of states (S_0, S_1, S_2, . . .) and rules for moving from one state to another. A state is defined as the set of all rules applicable to the state. Starting with any given state 0 there is a transition to a next state i via a symbol. Thus, for example, $S_0 \xrightarrow{a} S_i$ which can be written $(0, a, i)$; this is one rule of transition of the grammar. A rule of a grammar determines whether there is a direct transition from the present state to one or more other states, or to the same state itself in the form of a recursive loop.[2] Figure 1 is a graphical representation of the finite state generator used in the introductory Illustrative Study. The numbered circles represent the states of the system and the arrows represent the rules of transition permitted from one state to the next. From S_2, for example, the grammar permits us to move to S_5, then back again to S_2, thus: $S_2 \rightleftarrows S_5$. This loop can be traversed any number of times. The nonsense words associated with the arrows represent the finite set of symbols that comprise the given vocabulary of the language. This particular finite state system has a vocabulary of three words—NEN, PAG, and ZIR, which could be abbreviated as N, P, and Z for convenience. One of these words is generated each time we move from one state to the next. For example, in S_3 the rules say either move to S_4 by generating the word PAG, or make a recursive loop by returning to the same state itself by generating ZIR any number of times. Upon the first visit to $S_{0'}$ a sentence is completed; a sentence is any well-formed string. A well-formed string is generated by always starting in S_0 and following any route in the diagram and ending in the first visit to $S_{0'}$. The well-formed strings



RULES OF TRANSITION

(0,Z,1)	(1,P,3)	(3,Z,3)	(5,P,2)	(7,Z,8)
(0,Z,6)	(2,N,5)	(3,P,4)	(6,P,6)	(8,N,8)
(1,P,2)	(2,Z,4)	(4,N,0)	(6,N,7)	(8,P,0)

BASIC STRINGS	CANONICAL NOTATION
ZPPN	ZP (Z) PN
ZPZN	ZP (NP) ZN
ZNZP	Z(P)NZ(N)P

FIG. 1. A graphical representation of the finite state generator of sentences used in the Illustrative Study. (An admissible string is any sequence of words generated by starting in state 0 and ending on the first visit to 0′. All admissible strings can be expressed by means of the canonical notation given below the figure. Loops have been omitted in the basic strings. The rules of transition given beneath the diagram comprise an equivalent description of the generators. Each rule defines the state in which it is applicable, the symbol it generates, and the next state in which the rule leaves the system.)

[2] A recursive loop in a string is the unit that can occur in the string from zero to any number of times. Expressed graphically, a loop in the following string $S_0 \xrightarrow{a} S_1 \xrightarrow{b} S_2 \overset{a}{\curvearrowright} \xrightarrow{c} S_3$ occurs at S_2 and is indicated by the looping arrow a. A string could have any number of loops at any of the states. Each of the grammars in Experiments III and IV contains a double loop in one of its strings.

selected by the generator in Figure 1 are: ZPZN, ZPPN, ZNZP, ZPZPN, ZPNZP, ZNZNP, ZPNPZN, ZPNZNP, etc. Of these well-formed strings the first three we consider as basic strings. The important feature defining a basic string is that no state is visited more than once. When loops are inserted, states are revisited; basic strings with their loops we are naming formulas. Basic strings are significant because, as Chomsky and Miller have shown, they provide the pattern for all the longer sentences in the languages; in grammatical terms, they are the linguistic "frames" into which the elements must fit. Any number of different sentences can be generated by following the rules and using the recursive loops any number of times; if we indicate loops by parentheses, however, all the sentences in the system shown in Figure 1 can be represented by a canonical notation of the general formulae:

ZP(Z)PN
ZP(NP)ZN
Z(P)NZ(N)P

At the bottom of Figure 1 are stated the rules of transition, the basic strings, and the canonical notation for the formulas. For example, (0, Z, 1) is a rule of transition; it means that when the system is in S_0 it can generate the symbol Z and thus move to S_1. One direct path traced by a set of rules of transition, *without loops*, is a basic string: *with loops* it is a formula. Thus (0, Z, 1), (1, P, 2), (2, Z, 4), (4, N, 0′) yield the basic string ZPZN; by adding the following two rules, (2, N, 5) and (5, P, 2), a double loop has been added to the basic string which gives the formula ZP(NP)ZN.

Finite state systems provide attractive materials for experimental studies in psychology. Their formal simplicity and psychological complexity make them potentially useful for the study of higher mental processes. Moreover, the pattern and construction of finite state systems lend themselves to quantitative handling by information measures. These materials are a new, useful, and quantitative tool to apply to the study of memory and thinking. For these reasons we choose to use finite state systems as materials for exploring pattern conception. From a finite sample of sentences, subjects are expected to infer the underlying grammar. To what extent controlled variation in the formal structure of these systems may affect the behavior of the subject broadly defines the scope of the present research.

At the beginning of an experimental session a finite set of sentences from a particular language is presented to a subject in a rectangular array of cards on a table surface. The cards have been previously arranged by tables of random numbers and are all exposed simultaneously before the subject. (This procedure of presenting all the cards to the subject at once is different from the procedure followed in the Illustrative Study.) Each 3″ × 5″ card contains one string of words only; the string is typed in capital letters on a single line. The subject sits facing the cards; he reads the typed instructions, then repeats them to the experimenter in his own words.

Neutral instructions: "Here is an array of cards to be sorted into groups, into as many or as few groups as you wish to make and for whatever reason you wish to make them. When you are satisfied with your groups inform the experimenter."

If a subject has any questions regarding the time limit he is informed that there is none. Other questions are answered by merely referring him to parts of the typed instructions. When a subject begins to study the cards the experimenter starts a timer manually; the experimenter stops the timer as soon as the subject is through with his sorting. After the subject has made his groups he is asked to generate two sample strings to match each of his groups and the experimenter records the results of the subject's sortings along with any comments that the subject is willing to make about the way he performed the task.

The following principles were implicit in the procedure: (*a*) the problem is presented in as unstructured a way as possible in order to give the subject freedom to structure the materials by his own rules, (*b*) the subject is free of time pressures, and (*c*) there is a three-way check (sorting, additional instances, and free comments) on whether or not the subject was using any consistent strategy in arranging the cards into categories.

Subjects

Graduate and undergraduate students of Radcliffe College and Harvard University in the age group between 18 and 35 years, irrespective of field of specialization, were run as paid subjects. No subject served twice in any part of the experiment.

An Exploratory Study (Experiment II)

This experiment was a pilot study (*a*) to explore the kinds of strategies subjects use in grouping a given set of strings generated by finite state grammars (if, indeed, they used any strategy at all), and (*b*) to define some quantitative methods to measure the amount of "pattern" in each system of categories developed by a subject.

Strategies

Definition. A strategy may be qualitatively defined as the method or rule used by a subject for grouping a number of given strings into subsets. Operationally, a strategy is determined by the specific sentences that a subject puts into a subset. A specific strategy equates with a specific pattern a subject uses for making his groups. (This use of the term is different from the way Bruner defines it—Bruner et al., 1956, p. 54).

Quality. A strategy may be successful or less successful. It is considered successful if it corresponds completely with the basic strings of the experimenter's grammar. It is less successful if it deviates from the experimenter's rules. The degree of correspondence between the experimenter's rules and a subject's rules determines the degree of the subject's success. No strategy used by a subject is labeled as completely unsuccessful.

Procedure

The materials and procedure followed in this study have been described above; they are essentially the same in all the studies described below. In this exploratory study nine different finite state systems were set up in a 3 × 3 design; three different generators (grammar) of three, four, or five basic strings (see diagrams in Figures 2a, 2b, and 2c) were combined in all possible ways with three sizes of vocabulary, either of three, six, or nine words. Nonsense words were used in all the languages in order to minimize possible associations with the stimulus sequences. A tenth language consisted of sequences of random strings which were generated with the help of tables of random numbers. A vocabulary of three words was used for this set.

The languages were constructed with the following principles in mind: (*a*) Grammar III (five basic strings) is a hybrid composition from Grammars I and II. (*b*) The grammars are constructed to facilitate the experimenter in detecting objectively the strategy used by a subject. This principle places a constraint of congruence of certain parts—initial and/or final—in the structure of the basic strings. (*c*) The placement of symbols on transitional arrows is determined partly by the reason given previously, partly by haphazard assignment. (*d*) The number of instances (samples) in all ensembles in this study is 16. This particular design of three types of grammar and three sizes of vocabulary was used to include strategy construction under a variety of conditions to be tested more specifically later. The random strings condition was designed to serve as control for the nonrandom conditions. The nonrandom conditions are the nine grammar groups.

Sixty subjects were run individually, six subjects per condition. Subjects were distributed one per condition in the order in which they came.

Analysis of Strategies

Qualitative Analysis

There were 16 instances of sentences from each of the languages; the cards were numbered on the backs from 1 to 16. Each subject's sorting was recorded in terms of these numbers. The grouping produced by a subject was then compared with the various groupings that would

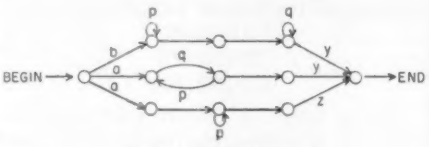


FIGURE 2a --GRAMMAR I

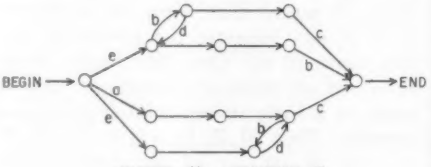


FIGURE 2b --GRAMMAR II

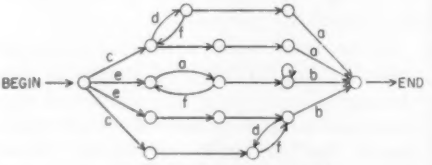


FIGURE 2c --GRAMMAR III

FIG. 2. Finite state generators for sentences used in the Exploratory Study (Experiment II). (The small letters on arrows indicate similarities and differences in the initial, final, and recursive-loop positions of symbols in the sentences generated).

result from following the different strategies. Thus, if the subject's groups were (3, 7, 15, 16) (1, 5, 10) (11, 12) (2, 6, 13, 14) (4) (8) (9), they corresponded with groups that resulted if the Length strategy were followed. Length strategy refers to grouping the instances by the number of words—the length of the sentence—contained on each card. The experimenter had worked out various groupings in terms of possible strategies so that a strategy could be objectively identified. Of course, the matching samples generated by a subject for each of his groups and the protocol were additional checks upon the identification of the subject's strategy.

Out of the billions of possible partitions of the 16 instances only about 20 strategies, not always markedly different from each other, were found to be used by subjects. The five most common strategies were labeled as follows:

Length: Subject puts together instances having the same number of words.

Alphabet: Subject alphabetizes the whole set of instances, treating it as a single category.

Initial: Subject sorts by similarity found in the beginnings of sentences. This similarity could range anywhere from one to four words.

Initial-Final: Subject sorts by the identicalness both of the beginnings and the endings of sentences.

Formula: This strategy approximates to the experimenter's rules for the finite state generator.

Miscellaneous: This category is a cover-all group for isolated cases of sorting such as sorting by a single striking word, or by the repetitive elements, or by rhythm, etc.

Quantitative Techniques

Since a variety of strategies was employed, strategies ranging from partially successful ones to the one that does succeed completely, it becomes necessary to order or scale the strategies. Earlier, we defined a successful strategy as one that completely corresponds with the basic strings of the experimenter's grammar. Three measures for scoring strategies were employed: an agreement-disagreement measure, a rating scale, and a time score. The first two were constructed specifically for evaluating strategies in this research.

Agreement-Disagreement Ratio (A-D ratio). This statistic is a kind of correlation technique for measuring the degree of agreement and disagreement between the experimenter's grammar and a subject's categories; or between any two subjects' categories. The ratio expressed (roughly) by the observed number of agreements over the total number of possible agreements in grouping, yields a measure of correspondence. When the two groupings are identical, the A-D ratio equals unity (see Appendix A for details of this technique).

Rating Scale (R scale). Each subject's strategy was assessed on a 10-point scale developed for this purpose. This scale was set up in the following way. Each subject's performance was classified into a strategy by the methods previously stated. Approximately 20 strategies seemed available to the subjects. Identical randomized lists of these were presented to five judges for rating. The judges were graduate students at Radcliffe College and Harvard University. They represented different fields of experience and cultural backgrounds. The judges were instructed jointly to rank on a 10-point scale the strategies subjects used. The instructions given to the judges were:

> You have been handed a list of labels for the strategies subjects used in an experiment in sorting into groups a pack of cards prepared by the experimenter. [Judges were shown the cards and the experiment was explained to them. Also, the labels of the strategies were described or explained.] You are being asked to rate these strategies by assigning them values between 1 and 10. Rate the poorest strategy 1 and the best one 10. You may assign the same value to several strategies if you think they rate the same and you may indicate distance in the quality of strategies by the size of the interval you leave between any two. Any questions?

After the questions were answered the judges worked independently. From the averages of the scores assigned by the judges to the strategies a standard 10-point rating scale for evaluating strategies was developed (see Appendix B for details).

Time Measure. A record of the total time required for sorting by each subject was made. Thus the number of seconds per word provided a third measure of a subject's performance.

Reliability of Measures

The A-D Ratio. Although this statistic is a fair measure of success in terms of the correspondence of a subject's responses to the experimenter's rules for generating the strings, it is subject to experimental artifact. For example, the Initial$_1$ strategy (instances grouped together on the basis of the first word only) yields a spuriously high correlation score. The number of subjects using the Initial$_1$ strategy is comparatively small, however, and the other two measures offer a second check on the results.

A second difficulty with this statistic is that interlanguage comparisons cannot be made without using some kind of correction factor for the number of instances in the set. The measure of agreement rises with the increase in the number of instances; the measure is not really comparable from one language system to another.

The R Scale. The following table shows the reliability of the ratings of strategies made by five judges. Rank-difference correlations were used.

It should be noted that the equality of intervals between ranks in this scale cannot be established.

TABLE 1

RANK-ORDER CORRELATIONS BETWEEN JUDGES FOR RATING STRATEGIES

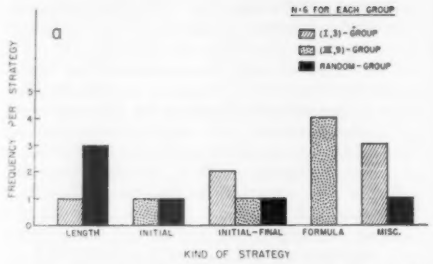
Judges	1	2	3	4
2	.89			
3	.94	.77		
4	.88	.89	.82	
5	.91	.87	.89	.88

Note.—From the ratings assigned by these judges to strategies was developed the 10-point rating scale.

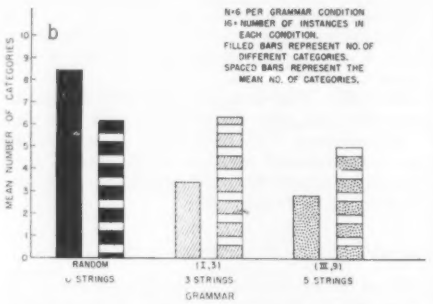
Time Score. Although time measures have been commonly used in studies of visual pattern perception (Adams, Fitts, Rappaport, & Weinstein, 1954; Anderson & Leonard, 1958; Fitts et al., 1956), in the present research they are only a rough measure. Several factors are simultaneously operative which decrease the reliability of the time scores. As the number of instances in the set increases, the time for sorting increases disproportionately, paralleling the rule that time increases according to the amount of information that must be processed. Both length and complexity of task tend to increase time, but if a subject works with the self-instruction of speed, he finishes the task much more quickly. The difficulty here is circular. The less time a subject spends with the problem the less information he extracts from the input and the less time he spends dealing with his problem.

Results

The distribution of subjects in the Random group is compared in Figure 3a with two of the nine groups (I,3 and III,9) with respect to the kinds of strategies used in each. The bars for the Random group and the (III,9) group (five basic strings and a vocabulary of nine words) are longer at the opposite ends. Condition (I,3) (three basic strings and a vocabulary of three words) has more affinity with the Random group than is apparent from the figure; three of its subjects who were classified under the Miscellaneous condition in the type of strategy used actually used some kind of differentiated counting technique that was slightly more complex than the plain counting involved in the Length strategy. Figure 3b compares the same three groups by the average number of different categories made by each in the



3a. Frequency comparison of strategies by subjects in the Random and two grammar groups. (The Random group worked with 16 strings that had been haphazardly generated. The (I,3) group worked with 16 well-formed strings that had been generated with three formulas and a vocabulary of three words. The (III,9) group worked with 16 well-formed strings that had been generated with five formulas and a vocabulary of nine different words.)



3b. Mean number of categories for subjects in the Random and two grammar groups. (Solid columns—first in each pair—compare the same three languages for the number of different categories made. Different categories refer to instance categories, e.g., one-instance category, two-instance category. Broken bars compare the same three groups for the average number of categories for each group.)

Fig. 3. Frequency comparison of strategies and mean number of categories.

first columns (filled bars). (By different categories we mean instance categories; for example, one-instance category, two-instance category, and so on.) Notice how the Random condition, where every instance belongs to a separate category, averages to the largest number of different categories. The second part (spaced bars)

compares the groups for the mean number of categories made by each. Here the difference between the Random group and (I,3) group is negligible.

Table 2 is a frequency tabulation of all subjects showing their distribution by the kinds of strategy they employed and the experimental conditions under which they had worked. Note that there are just about five or six types of strategy that cover all subjects under the various conditions tested. Also note that the better or the more successful strategies are used under some conditions more than in some others. For example, none of the 18 subjects used the Length or the Alphabet strategies when the number of alternative words used was nine. Ten of these subjects used some form of the Formula strategy. When the sentences were formed with a vocabulary of only three words, on the other hand, 5 of the 18 subjects used the simpler strategies and only 4 discovered the Formula strategy; all of those 4 subjects came from the same condition group.

For each of the 10 experimental conditions, Table 3 presents means which have been calculated by assigning values to strategies by the statistical techniques discussed above. Because the experimenter worked with no rules in the Random set no A-D ratios can be calculated for that strategy. Ratios for Condition III are found to be higher for each constant condition of vocabulary size, except in the grammar Group (II,3), which may be treated as an exception in all the tables given subsequently because of an artifact in the construction of the language. The group means for the R scores are tabulated in the second row. Here the Random group and (I,3) group are more alike, as we have com-

TABLE 2

FREQUENCY DISTRIBUTION OF SUBJECTS IN THE RESPECTIVE LANGUAGE GROUPS BY THE TYPE OF STRATEGY THEY USED

Grammar	Kind of Strategy						*N*
	Length	Alphabet	Initial	Init.-Final	Formula	Misc.	
Random$_3$	3		1	1		1	6
(I,3)	1			2		3	6
(II,3)	1	1			4[a]		
(III,3)	1	1	1	3			
Total	(3)	(2)	(1)	(5)	(4)[a]	(3)	18
(I,6)	2		1	1	1	1	
(II,6)	2		2		2		
(III,6)			2	2	1	1	
Total	(4)	—	(5)	(3)	(4)	(2)	18
(I,9)				2	4		
(II,9)			3	1	2		
(III,9)			1	1	4		
Total	—	—	(4)	(4)	(10)		18
Sum Total	10	2	11	13	18	6	60

Note.—Figures in parentheses are sums of frequencies immediately preceding those totals.
[a] These four subjects all belong to the (II,3) group.

TABLE 3
GROUP MEANS FOR TEN EXPERIMENTAL CONDITIONS

Mean Score	Language Group									
	0,3	I,3	II,3	III,3	I,6	II,6	III,6	I,9	II,9	III,9
A-D ratios	—	.46	.773	.616	.437	.601	.73	.725	.805	.85
R score	1.83	2.83	7.0	3.71	3.66	4.66	3.33	6.83	5.5	7.2
Time score	6.04	6.52	4.04	6.36	3.61	4.75	4.64	4.09	.41	6.05

Note.—The means have been calculated by three independent statistical measures: A-D ratios, Rating scores, and Time scores. Time scores represent average time taken per syllable.

mented before, and both are inferior to Groups (I,9), (II,9), and (III,9). The time comparisons in the third row are somewhat fuzzy. Sometimes a subject, who may have worked with the instances for a very long time, will settle on a very poor strategy at the end; sometimes, the opposite is true.

A comparison of the group means by vocabulary and by grammar by applying all three measures is made in Table 4. Although the figures given in this table are not always directly comparable, their relative increase or decrease with change in the experimental condition is constant.

Discussion

This first study was intended to explore the kinds of responses subjects make to a categorizing task. The task as we have defined it is one of sorting strings of words simultaneously presented to a subject on small cards. The subject is free to work on these instances for as long as he wishes, without receiving further help from the experimenter, however, or from the instructions that are given to him. He is neither told how many categories he should make, nor is he given any indication of success or failure. The criterion of success, if any, is set by the subject himself and he does not have to form groups if he thinks no groups can be made.

Though the experimenter is somewhat in the dark as to what the nature of the controlling variables is, he does, however, anticipate a few problems. Therefore, he sets up his first study under a variety of experimental conditions to get leads about the possible functional relations between some of the experimental variables and subjects' responses. The basic variables that the experimenter works with are: the number of formulas which provide the

TABLE 4
A COMPARISON OF RELATIVE DIFFICULTY OF GRAMMAR AND OF VOCABULARY AS INDICATED BY THREE INDEPENDENT MEASURES

Grammar:	Random	I	II	III
A-D ratios	—	.53	.72	.73
R score	1.83	4.44	5.7	5.1
Time/Word	6.04	4.85	5.91	5.59
Vocabulary:	**Random**	**3**	**6**	**9**
A-D ratios	—	.62	.59	.79
R score	1.83	4.6	3.9	6.6
Time/Word	6.04	5.48	4.51	6.24

basic framework for the strings and the number of different alternative words with which he develops his formulas.

The results illustrate several important features about subjects' responses in an information processing situation. First, subjects process information by rules, or as we have suggested earlier, by pattern conception. Next, we note that the rules formulated by more than 90% of subjects can be classified under five headings: Length, Alphabet, Initial, Initial-Final, and Formula strategies. Third, there is a hierarchy within these five or six types of strategies which has been identified in subjects' reponses. These strategies vary from simple to complex. Finally, the results indicate that some experimental conditions are better than others for inducing the formulation of the more successful strategies. The reader will recall that the more successful strategies have been defined as those that have a higher agreement with the experimenter's formulas. The Random strings, which were haphazardly constructed, embody a condition which is least favorable to the emergence or acceptance of pattern strategies in subjects' behavior. Evidence from our data suggested that vocabulary and grammar may be two of the more effective parameters in pattern conception.

Pattern Conception. What we had observed as a dramatic instance of orderly responses demonstrated in the introductory study, was confirmed by the responses of more than 50 subjects in this exploratory experiment. It is not a fortuitous matter that no subject, not even one from the Random group, handed the cards back to the experimenter in some haphazard arrangement or declared the task insoluble. On the contrary, in each case the subject used some kind of a system by which he had looked for relations among the instances to create his categories. Even with the Random group some subjects worked very hard and long trying to find some sort of scheme to organize their information. By and large, the subject could verbalize the type of relations that he had discovered in terms of a rule or rules though he was not always able to do this with precision.

Hierarchy in Rules. The thing that amazed us most was that nearly all subjects could be classified so easily within the same general types of five or six strategies—strategies which were, by and large, the same ones we noted earlier in the responses of our single subject described above in the Illustrative Study. In other words, a hierarchy in rules that was indicated in the individual's responses seems to characterize the group responses as well. This is a two-sided observation: it points first, to the fact that responses under all conditions of card sorting described above are classifiable within five or six easily detectable strategies; and second, that these strategies themselves can be ordered in complexity.

We can classify these strategies into two types: the ready-made type and the invented type. Ready-made rules are independent of the nature of the information to be processed. The rule of counting, for instance, to make groups of equal length, already exists in a subject's repertory and can be easily transferred to almost any kind of material. The ready-made rule applies whether the materials be heads of cattle, pictures in an art gallery, or the coded messages about the enemy's moves. This rule of counting is an instance of a simple quantitative strategy that is highly useful in dealing with certain kinds of information, but not so appropriate for certain other kinds. The counting rule may take on slightly different shades, as counting the number of different words, or counting the number of repetitions of any word, etc., but essentially it is the same rule that is applied.

Alphabetizing is here classified as a ready-made strategy because this rule, too, already exists in the everyday repertoire of a college student. Since it is always possible to alphabetize a given series of words of any language, it requires little ingenuity on the part of a subject to apply this rule.

It is important to note that a subject loses information by applying the simple strategies of counting and alphabetizing to these not-so-simple sequences. We have.

therefore, evaluated such strategies as being less successful. A strategy is rated as less successful in proportion to the amount it deviates from the finite state grammar which governs the given strings. In the simpler strategies a subject seems to work more with the perceptual characteristics of the information than with the structural ones.

Subjects in the Random group, which is the control group for all the nine experimental groups, seem to resort to the simple strategies, particularly the one of counting, more than the subjects in the grammar groups seem to do, except the subjects of the (I,3) group. The responses of this group have close kinship with the responses of the Random group. When a subject in any group either finds the task too difficult for him or when he does not wish to exert himself, functionally, he is treating the grammar strings as if they were random strings. Hence the strategies that subjects would be found to be using in these instances would be close kins to counting or to alphabetizing strategies. It seems highly probable that subjects in the (I,3) group were treating the grammar strings as random strings.

The observed fact that a subject in the Random group follows a counting strategy does not necessarily imply that the subject has worked with one strategy only. Subjects' protocols bear evidence to the contrary. Several subjects described complex strategies that they had tried to apply to the strings, but which they had discarded because they always seemed to find too large a number of single cases. Notice that a subject is uneasy about having too many one-case categories, and this is virtually what he would find in the random set where each instance is unique. After having searched for abstract relations, a subject is forced by the stimulus properties of the set to settle for something less. This sample of behavior is apropos to the basic argument in this thesis.

Strategies labeled Invented are different from the ones just described. We classify the Initial_n-Final_n strategy and the Formula strategy under this label because it is the nature of the materials which suggests these strategies to the observer. A high degree of discovery, however, is also involved in these strategies, but since a subject is unaware of the embeddedness of any formulas within the materials, we prefer to call them the Invented rules rather than the Discovered rules. This label, however, applies most appropriately to the Formula strategy. In this strategy the subject has to go beyond the obvious information. The subject has to tease out an invariant sequence of words in each of the instances, as also he must find the recursive units. The search for invariance and variance yields formulas which are the basic framework for the grammar strings. This type of response calls for a certain degree of abstraction in order to discover the true pattern implicit in the strings.

Invented strategies correlate more closely with the information given. Since the instructions a subject receives are neutral, his only source of information is the set of strings he has before him. The amount of information he is able to extract from them is a critical factor in the way he will process the information. His efficiency in extracting the maximum amount of useful information depends largely upon two, sometimes interdependent, conditions of motivation and stimulus complexity.

Motivation. Each subject seems to redefine the problem to himself and it is very important how he does define the task. Subjects who look for quick and easy solution—whose motivation appears to be to escape the situation as quickly as they can—do not get themselves involved in the real problem of finding relations among sets of instances. They tend to manipulate the more obvious characteristics of the stimuli and fail to extract useful information contained in the sequence order of the words. Though they might also see other little bits of isolated information, they tend to ignore these and proceed to apply the easy rules of counting and alphabetizing, thus minimizing the cognitive strain.

Stimulus Complexity. Subjects who do not get seriously involved in the task may

or may not end with the Invented strategies. One subject in the Random group worked with her 16 cards for 41 minutes. She tried to make groups by initial similarities, by ratio relations, by sequence relations, and by a method of substitution. But when none of these strategies worked by her criterion of success, she resorted, according to her own version, to "the simplest strategy of counting." The task of seeing relations among the random strings is perhaps psychologically more complex than the task of seeing relations where relations do in fact exist.

The results from all nine grammar group conditions of finite state systems where relations do in fact exist, do not distribute under the Invented strategies only. Some do and some do not. In Table 2 we notice that the stimulus condition, where nine alternative words were used to develop the strings of sequences, as compared to three in the Random group, and three and six in the other two conditions, more than 50% of the subjects evolved the Formula strategy as compared to the 0% of the Random group and 22% in each of the other two. Of these, the 20% of subjects in the vocabulary three group, all come from a single condition, (II,3), meaning that none of the subjects from Condition (I,3) or (III,3) or (0,3) arrived at the Formula strategy.

Are these differences, indicated in the percentages of subjects who evolve the Invented strategies and those who do not, functionally related to the complexity of the experimental variables under which subjects work? If so, what constitutes experimental complexity? That is the question we turn to in the next section.

Experimental Variables

In the exploratory study of the previous section evidence for a search for pattern was found. Sometimes this search was creative, sometimes transitive, but what was not clear from the data was why this difference existed in the quality of performance. Informal evidence suggested that instruction was one variable that needed to be controlled more precisely and the data indicated that some stimulus features of the input were contributing to the outcome. Accordingly, some experiments were carried out to assess the effect of the following four variables:

1. Instructions. Four different instructions were tested to select one that would maximally involve the subject in the task, but with minimal guidance from the nature of the instructions.

2. Number of instances. The number of instances per formula was varied from two instances to three to four, to test the effect of the amount of evidence on the quality of strategy.

3. Number of different alternatives (Vocabulary). Three or nine different words comprised the vocabulary for developing the different language systems.

4. Structural similarity (Grammar). The effect of the formal structure of the languages, three rules as against five rules with variation in similarity between the rules, was also investigated.

The Instructions
(Experiment III)

Materials

Two new artificial languages, $(I,9,16)_b$ and $(III,3,23)_e$ were constructed to study the effects of four types of instruction on each. Language $(I,9,16)_b$ is constructed with three formulas in its grammar; its vocabulary consists of nine words. Sixteen instances were selected to make up this set. Language $(III,3,23)_e$ contains five formulas; three words have been used for its vocabulary. This set contains 23 instances of well-formed strings. Two additional language conditions, (I,3,16) and (I,3,23), were also tested, both under "eliminate speed" instructions. Altogether, then, there were 10 language conditions that were compared in this study.

Procedure

The same general procedure described earlier was used in all experiments reported in this section. The specific variable to be considered here is the kind of instruction to be used. Ten groups of subjects, four subjects per condition, were run under four different instructions. No subject served twice in any part of the experiment. Subjects were distributed, one per condition, in the

order in which they came. The following four sets of instructions were used:

1. Neutral. "Here is an array of cards to be sorted into groups, into as many or as few groups as you wish to make, and for whatever reason you wish to make them. When you are satisfied with your groups inform the experimenter."

2. Eliminate Speed. "Here is an array of cards to be sorted into groups, into as many or as few groups as you wish to make, and for whatever reason you wish to make them. Your performance will be judged *not* by speed, that is, not how quickly you are through, but by *what* you do. When you are satisfied with your groups inform the experimenter."

3. Relations. "Here is an array of cards to be sorted into groups, into as many or as few groups as you wish to make on the basis of some relations that you can find in the members of the groups. Your performance will be judged *not* by speed, that is, not by how quickly you are through, but by how well you discover relationships. When you are satisfied with your groups inform the experimenter."

4. Logical. "Here is an array of cards to be sorted into groups, into as many or as few groups as you wish to make on some logical basis. Your performance will be judged *not* by speed, that is, not by how quickly you are through, but by how logical you are. When you are satisfied with your groups inform the experimenter."

Results

Table 5 presents the means, ranges, and critical values for the A-D ratios, R scores, and Time scores for 40 different subjects run under 10 different experimental conditions. The critical differences between groups were calculated by the Link-Wallace method of short-cut allowance for one-way analysis of variance (Mosteller & Bush, 1954, p. 305). This statistic uses the range as a measure of variation instead of the sum of squares.

Conditions F, G, H, and I in Table 5 apparently belong together. The difference between any two totals in this group is below 1.51 in the A-D ratio comparisons, below 18.80 in the R score comparisons, and below 41.59 in the time comparisons. It is clear, then, that variation in instructions fails to produce any recognizable difference in the performance of subjects in language $(III,3,23)_c$. Presumably, the task involved in this language is too hard for the subjects. Even when the subjects (Column H) took longer than subjects in any other group, they failed to give efficient results. Since the complexity of this language masks the effect of instruction (if any), we shall concentrate on the simpler Grammar I.

Conditions C, D, E, and J likewise obviously belong together. None of the differences between any two of these con-

TABLE 5

MEANS, RANGES, AND CRITICAL VALUES

	Language	Instructions	A-D ratios			R score			Time score		
			M	*R*	*CV*	*M*	*R*	*CV*	*M*	*R*	*CV*
A	(I, 3, 16)	Elim. Speed	.333	.40		1.75	3		2.38	1.39	
B	$(I, 9, 16)_b$	Neutral	.402	.433	1.95	4.25	7	0.47	3.55	3.56	41.58
C		Elim. Speed	.69	.25		7.5	6		3.37	2.25	
D		Relations	.95	.20		9.75	1		5.49	4.35	
E		Logical	.76	.14		7.5	6		5.35	2.71	
F	$(III, 3, 23)_c$	Neutral	.58	.21	1.51	4.0	2	18.80	7.81	16.94	41.59
G		Elim. Speed	.50	.43		5.25	5		9.04	5.99	
H		Relations	.58	.27		5.0	4		18.58	22.59	
I		Logical	.49	.43		5.75	5		8.55	17.70	
J	(I, 3, 23)	Elim. Speed	.79	.46		8.75	1		9.76	10.09	

ditions exceeds the calculated critical values in any of the three measures. All of these four conditions belong under the simpler Grammar I, which also includes Conditions A and B. Subjects who served under Condition A performed most poorly, even though, as in C, D, E, and J, they were instructed not to speed. Subjects for $(I,9,16)_b$ who served under Neutral instructions (see Column B) do not resemble or differ consistently with the other groups.

Discussion

In the preceding section we raised a question about why some subjects follow easy strategies and some others follow more complex ones. An analysis of the subjects' protocols gave an indication that one necessary, though not sufficient, condition for evolving a good strategy is that the subject should extract all the *useful* information from the instances he receives. This business of extracting the most useful information is in itself a very critical step in pattern conception. The success with which a subject decodes relevant information from the stimuli depends partially upon the degree to which he involves himself in the task. This in turn is influenced by a variety of factors, such as the past history of the subject, his present motivation, his age, etc., as well as the instruction he receives from the experimenter regarding the nature of the problem, other things being equal. An individual's history or his motivation is difficult to control by laboratory techniques, but the experimenter can explore the instruction variable.

Actually, a subject receives two kinds of information from the experimenter. One kind comprises the main body of information that is contained in the strings of nonsense words with which the subject is supposed to work; the other is some information *about* this main body of information to be categorized. The main body of information is contained in what Bartlett (1958) has called a "closed system," which the subject must discover. The efficiency with which a subject processes the main body of information will in part depend on how well he has decoded the information from the instances presented to him. It is the experimenter's wish to get a subject started on this search but without telling him how to proceed. Thus in setting up the instructions the experimenter keeps two things in mind: to get a subject maximally involved in solving the problem and to do this with a minimum of guidance from outside.

Neutral instructions are suspect when they are evaluated by this two-edged criterion. In the absence of specificity, both of instruction and of materials (nonsense strings), a subject may often resort to self-instruction, either to work for speed, or to find some easy-way-out strategy; he thereby loses much information. To overcome this difficulty it seems necessary to furnish a subject with some other kind of instruction.

French (1954a) has shown that performance improves with more specific instructions in a task that involves sequential dependencies. The question here, however, is not one of improving a subject's performance by giving him more guidance or by making the instructions more specific—we take that for granted—but rather the question is one of finding the kind of instructions that will induce a subject to make the maximum effort to solve the problem, and to bring this about with as unstructured a kind of working instruction as possible.

We noted in the results that when a language is too difficult for the subjects, as $(III,3,23)_e$ seemed to be, the effect of instructions is confounded. But with the simpler Grammar I, subjects do better when they are instructed to slow down, with the one notable exception of (I,3,16) —see Condition A. We pointed out earlier that the results of Condition A are the poorest of all the conditions tested. A differs from B, C, D, and E in the size of the vocabulary; it differs from F, G, H, and I in its grammatical structure; and its difference from Condition J seems to be clearly related to the number of instances. Presumably, A suffers in part from one or more of these conditions.

The difference between B and Conditions C, D, E, and J was attributed to different instructions. All four conditions, C, D, E, and J, required subjects not to rush (note that Relations and Logical also eliminate speed), and they all emphasized that it is *what* a subject does that is more important than speed. Condition B is neutral to these emphases. It does not seem unreasonable, therefore, to suppose that the emphasis on "what you do" increases the search for useful information, and this search for the useful information, we have granted, is one critical step in pattern conception.

Johnson (1955) discusses the effect of instruction on performance under the concept of "set" (see pp. 64-67). After defining set as "temporary integration of behavior," he elaborates on its meaning as integration in terms of the obvious and simple things, or integration in terms of "subtle aspects of complex stimulus patterns." Explained in this way, set seems to be no different from recoding of information. Verbal instructions can drastically change the quality of recoding. As for instance, in the Carmichael, Hogan, and Walter (1932) experiment the same figures are recoded as dumbbells by one group and as spectacles by another group of subjects, depending upon the instructions the group received. In studies such as this one, the experimenter has already laid down the recoding scheme for his subjects through the instructions he provides, and then he is surprised (and pleased perhaps!) to find that subjects have followed his scheme.

It is this kind of coaching in the recoding scheme of a subject that we wished to avoid in our instruction. For, after all, a major interest in the present research is to study those factors that affect the recoding scheme of a subject. Very little new information will be gained by recovering the same coins that we originally fed into our slot machine.

So, what do we finally conclude? On the whole one could say that the instructions, to the extent to which they were varied, did not produce a strong effect on the quality of the performance. Other variables (complexity of grammar, number of different alternatives, number of instances) seem to be more important because they can make the problem so hard that they mask out the effects of instructions. When conditions are optimal for instructions to have an effect, the main effect that changed instructions produced was to slow subjects down to make them stop, look, and think more carefully, and this is the only effect the experimenter wished to induce through instructions. Any more help than just maximally motivating subjects to think would be a violation of what has been described as good instruction. Since the instruction to eliminate speed, together with the emphasis on "what you do," is common to all three sets of instructions tested, it seems reasonable to suppose that this amount of change in the neutral instructions was enough to involve the subject in the task. Therefore, we adopted Eliminate Speed instructions as standard for all subsequent studies.

The Number of Instances (Experiment IV)

Materials and Procedure

One important characteristic of an ensemble of instances is its size. The size of the set can be controlled by controlling the number of formulas of a grammar and/or the number of instances given for each formula. Moreover, one might like to control the total number of words in the ensemble. The structure of the languages used in this study is clarified in Table 6.

The Random condition contained no formal grammar; it served as the control group for all the other conditions. Languages B, C, and D are identical except for an orderly increase from two instances per loop in (I,3,13) to three instances in (I,3,18) to four in (I,3,23). The number of instances per loop and the size of the set are interrelated; as instances increase in number, the size of the set increases also.

Languages (I,3,13) and (I,9,13) are also identical except for the difference in the size of the vocabulary: B is developed with three words; E with nine. Thirteen instances for Grammar I have been considered to comprise a minimum set by definition. Condition F, $(I,9,16)_b$, represents a random increase in the number of instances from 13 to 16. The choice of the size, 16, was made to match the size of the Random set. Conditions E and F will be compared for a small

TABLE 6

STRUCTURAL ANALYSIS OF LANGUAGES

Language	Grammar or number of formulas	Vocabulary or number of different words	Instances per loop	Size of set	Total number of words
A Random (0,3,13)	0	3	1	16	118
B (I,3,13)	3	3	2	13	90
C (I,3,18)	3[a]	3	3	18	180
D (I,3,23)	3[a]	3	4	23	180
E (I,9,13)	3[a]	9	2	13	90
F $(I,9,16)_b$	3[a]	9	Irregular	16	118
G $(III,3,23)_c$	5	3	2	23	180
H $(III,9,23)_c$	5[b]	9	2	23	180

[a] Same as in B.
[b] Same as in G.

increase in the number of instances from 13 to 16, other things being held constant.

A similar comparison in size is envisaged for Languages $(III,3,23)_c$ and $(III,9,23)_c$. Grammar III consists of five formulas as against three of Grammar I. A minimum set for Grammar III consists of 23 instances, which also allows 2 instances per loop. All experimental conditions for these two sets were held constant, except for the size of their vocabulary.

One could also compare (I,3,13) and $(III,3,23)_c$ with each other as minimum sets for their respective grammars. Since the number of instances per loop has been held constant, any differences found between the two languages would be attributable either to the number of formulas in the grammar or to the number of instances in the sets, or to both.

Ten subjects per condition, using the instructions aimed to eliminate speed, were run for each of the eight language conditions.

Results

Table 7 presents the means and critical values for 10 language conditions by the usual three measures. Only 8 of these 10 groups, A through H, concern us in this section. We note first of all that in Grammar I, Conditions B, C, and D, as the instances per loop increase from two to three to four, strategies improve correspondingly. In working with 13 instances a subject spends about the same amount of average time per word as a subject working with the larger sets of 18 or 23 instances, but he does not achieve better results. In fact, the strategies used under this condition are poorer than most and about as poor as the ones used by the Random group, which used the poorest strategies.

TABLE 7

MEANS, RANGES, AND CRITICAL VALUES

Language Group	A-D ratio	R score		Time score	
	M	*M*	*R*	*M*	*R*
A Random (0, 3, 16)	—	2.8	6	11.38	49.24
B (I, 3, 13)	.253	3.1	7	8.14	29.93
C (I, 3, 18)	.389	4.6	7	6.71	23.02
D (I, 3, 23)	.598	6.5	8	7.73	14.27
E (I, 9, 13)	.532	6.7	6	4.996	8.29
F $(I, 9, 16)_b$	.748	8.2	7	3.916	4.23
G $(III, 3, 23)_c$	.419	4.7	8	5.396	7.31
H $(III, 9, 23)_c$	.706	6.0	7	5.035	6.53
I $(III, 3, 23)_d$	.541	3.5	8	4.998	11.05
J $(III, 9, 23)_d$	.81	7.1	7	5.478	14.39
Critical Value		33.37		79.08	

Next we note that without any increase whatsoever in the number of instances, if we merely changed the size of the vocabulary from three words to nine, performance is drastically modified (cf. Column B with E and G with H). Now look at the time measures. Subjects in Column F take far less time than subjects in Column B and yet they do far better than subjects in Column B. The same is again true, though not quite to the same degree, of subjects in $(III,3,23)_c$ and $(III,9,23)_c$. Apparently, one must look for some other factor(s), instead of the number of instances, to account for these differences between the groups.

A further point of interest may be noted from a comparison of Languages (I,9,13) and $(I,9,16)_b$. Earlier we explained that the two languages are identical in every respect except for a fortuitous increase in the number of instances from 13 in Condition E to 16 instances in F; the results, however, that follow this change are not so fortuitous. Even a small random increase in the number of instances in a language with the larger vocabulary apparently makes some difference in the subjective redundancy of the set for the subjects. But when we compare Columns A and F, both sets with 16 instances in each, the Random set scores are much inferior.

Now compare the minimum sets for both grammars. We consider 13 instances for Grammar I and 23 instances for Grammar III, as minimum sets for each, by definition. (I,3,13) is not too different from $(III,3,23)_c$ nor is (I,9,13) very different from $(III,9,23)_c$. Presumably, the effect of the size of the set is either negligible or it is masked by the effects of other variables.

To summarize our observations, three points seem rather clear: the larger vocabulary decreases difficulty, more instances per formula also decrease difficulty, the effect of the total number of instances in a set is easily modulated by other variables.

It is not clear, however, as to what effect an increase in the number of formulas produces upon subjective difficulty.

Discussion

Two of the four points summarized above pertain to the number of instances in the set; they are not unrelated points. The total number of instances in a set is automatically increased with each new addition of an instance for any of its formulas. This means that two things happen for a subject. First, every single addition of an instance to the set increases the difficulty of the set by increasing the number of relations the subject must now study (if n cards have m relations, then $n+1$ cards will have $m+n$ relations). Second, every addition of an instance for any of the formulas contained in the set simultaneously reduces difficulty by increasing the amount of evidence available. This latter generalization, however, is not true for a random set of strings; an increase in the number of instances in the Random set does not reduce difficulty, but rather increases confusion.

In the context of the present discussion we define redundant strings as those that have been derived from the same grammatical formulas. They repeat information by replicating formulas or instances or both. Random strings, on the other hand, contain no formulas or any orderly repetition of information. They could be viewed as containing as many formulas as there are instances in the set, as was suggested previously; as such, they will contain only one instance per formula. A redundant set which has been generated with one, two, or more formulas will contain instances which repeat information about the formulas. Thus the number of instances per formula and the number of formulas that constitute the grammar are somewhat interdependent features in a language.

Random strings are confusing to a subject because they are difficult to package. They contain more information than any redundant language of comparable size contains (see Columns A and F) and this mass of disjointed information is difficult for a subject to handle. The larger the size of the Random set the more difficult it is

for a subject to process the information it contains. Thus, the critical factor for increasing the efficiency of performance is not one of increasing the amount of haphazard information, but it is one of receiving a greater amount of more *useful* information. Reiss (1958) came to the same conclusion with his subjects when he varied the amount of initial information presented.

In the absence of repetition (redundancy) of information, relations between instances are few and far between. The subject, who is usually disposed to sorting his environment by searching for redundancy, is nonplused in a random set-up. Now introduce some redundancy into the situation by presenting to the subject two instances per formula, instead of one instance or none, and performance improves. Bartlett (1958) made this same observation on his subjects in tasks that required gap-filling by methods of interpolation and extrapolation. In the case of a simple formula, a string with one loop only, if one now gives the subject just one more instance, making it three instances per loop, the improvement is phenomenal. The subject's response almost begins to approximate to the grammatical structure of the language. One is led to think that depending upon the amount of complexity of the formulas, one could expect an upper ceiling in improvement in performance beyond which an increase in the number of instances ceases to be effective.

When the number of instances per loop is small, say zero or one, it is more possible to confuse the instances of one formula with the instances of another on the basis of superficial relations—what Wertheimer (1945) calls "and-summative aggregations" (p. 190). Actually in the ideal situation, where instances have been grouped according to the formulas they exemplify, instances which are positive for one formula are negative for all other formulas in the set. This, however, is true only when a subject has clearly seen the structure of the formulas in abstraction. Pattern conception requires the observer to make a two-fold response to the stimulus ensemble; a response of the ingroup type and of the outgroup type. The positive instances for a formula the subject must "aggregate" as members of the ingroup for that subset; all other instances become negative instances for this formula and which he must "segregate" as the outgroup members belonging somewhere else. In effect, the subject is making an inductive-deductive response. Hypothesis making and hypothesis testing are going on simultaneously as a subject decides which instances are positive and which are negative for the particular category he is working with. His certainty increases as the number of instances that fit his hypothesis increases. More instances per formula increase evidence and consequently decrease the probability of error. It is a well-known fact in communication engineering that one repeats a message as a safety device against "noise" in the channel. The noise which the negative instances create for a subject in our channel can be reduced by increasing the number of positive instances, other things being equal.

Thus we are led to conclude that within certain limits, the number of instances per formula reduces difficulty, subject to certain interaction effects from other variables. Although an increase in the size of the ensemble, within a reasonable range, increases the difficulty of the set, the quantitative aspect of the set is relatively unimportant. If, however, the size of the set is too large the problem is either rendered insoluble, or it becomes ridiculous, depending upon the number of basic strings in the set.

The interaction effects we referred to above have pointed in the direction of the size of the vocabulary and of the complexity of grammar as two possible variables that affect pattern conception. In fact, one of the points we noted from our results was that a large number of different words reduced the difficulty of the task for the subject. We could not see our way clearly on the question of grammatical complexity. Let us explore the effects of these two variables further.

The Number of Different Alternatives and Structural Similarity
(Experiment V)

Two variables that have been clouding the analysis of our data are the parameters of the number of different words and of structural similarity of the strings. We have equated the number of different alternatives with the vocabulary of a language, and its grammar, roughly, with structural similarity. Two sets of alternatives, consisting of three words and of nine words, constitute the size of the vocabulary of the languages to be tested. The generators for these languages consist of three basic strings, five basic strings with replication and five basic strings without replication. The Random set of strings has no grammar, of course.

Materials

Two additional languages, (III,d), with vocabularies of three and nine words, respectively, were constructed to control similarity in the formal structure of the grammars. By similarity in structure (in this context) is meant: (*a*) One grammar contains some information about the other grammars. For instance, Grammar (III,d) contains five formulas, three of which are identical with the three formulas of Grammar I, and with three of the five formulas of (III,c). (*b*) Structural similarity also means that one formula contains some information about the other formulas. For example, Formulas 4 and 5 in Grammar (III,c) repeat Formulas 2 and 3; in (III,d) they do not. This explanation of the structural

TABLE 8

t VALUES FOR A-D RATIOS AND R SCORES

		A-D ratio									
R score		Random A	(I,3,13) B	(I,3,18) C	(I,3,23) D	(I,9,13) E	(I,9,16)$_b$ F	(III,3,23)$_a$ G	(III,9,23)$_c$ H	(III,3,23)$_d$ I	(III,9,23)$_d$ J
Random	A										
(I,3,13)	B	.03		1.97	3.35	3.00	7.35	.84	5.27	3.69	7.15
(I,3,18)	C	1.75	1.15		2.09	1.60	5.78	.16	3.88	2.41	5.77
(I,3,23)	D	2.23	2.76	1.40		.56	1.53	.90	.96	.57	2.00
(I,9,13)	E	3.93	3.43	1.8	.15		2.45	.58	1.69	.10	2.88
(I,9,16)$_b$	F	6.28	5.93	3.49	1.45	1.51		1.79	.52	3.35	.86
(III,3,23)$_a$	G	2.02	1.55	.04	1.42	1.90	3.60		1.51	.66	2.06
(III,9,23)$_c$	H	2.66	2.50	1.10	.36	.57	1.99	1.09		3.25	1.16
(III,3,23)$_d$	I	.07	.03	1.90	2.35	2.89	4.74	1.10	2.08		3.68
(III,9,23)$_d$	J	3.61	3.30	.15	.42	.31	.94	1.90	.81	2.85	

Note.—*t* values for A-D ratios are entered in upper triangle. R score *t* values are entered in left triangle. Since A-D ratios were not directly comparable for different sized sets, A-D *t* values apply for sets of equal size.

p =	.1	.05	.02	.01	.001	df = 18
t =	1.734	2.101	2.552	2.878	3.922	

similarity in the grammars can be schematized thus:

Grammar I	Grammar (III,c)	Grammar (III,d)
1. - (-) - - -	1. - (-) - - -	1. - (-) - - -
2. - - (- -) - -	2. - - (- -) - -	2. - - (- -) - -
3. - (-) - - (-) -	3. - (-) - - (-) -	3. - (-) - - (-) -
	4. - - (- -) - -	4. - - - ((-)(-)) -
	5. - (-) - - (-) -	5. - (-) - (-) - (-) -

Results

In Table 8 are shown results of the *t* tests computed to evaluate the significance of the differences between pairs of conditions. It can be seen from this table that performance is significantly better with a vocabulary of nine words than of three (cf. B with E, G with H, and I with J). But when the grammar changes from three basic strings to five, R scores remain pretty much the same.

The A-D ratios rise automatically with an increase in the size of the set, as we noted earlier, so that the differences do not reflect a difference in the subjects' performance. Time, except for (I,3) remains constant for the several conditions.

In comparing the minimum sets of (I,3,13) and (I,9,13) with each other, we are impressed with the difference between the two conditions as evident from all three measures in Table 5. The reader will recall that these sets are identical in every respect except for vocabulary; the number of different words, then, does affect the nature of the strategy used. But when we turn to Language (I,9,13) and (I,3,23) which vary in two parameters—in vocabulary and in the number of instances per formula—we notice that the R score differences vanish, that the A-D ratios for (I,3,23) should be (and are) higher by virtue of the size of the set, and that time differences now appear.

The Random group and the (I,3,13) group arrive at similar strategies but at divergent speeds (cf. Columns A and B in Table 8).

Table 9 compares two types of strategies in terms of speed of sorting. Individual data have been clipped for the two extremes of the strategies rated as low and high. Column A is a listing of the amount of time spent per word by subjects who were rated low, that is, who scored either 1 or 2 on their strategies, and Column B presents the data for all subjects who scored 9 or 10. We remind the reader that

TABLE 9

TIME SPENT PER SYLLABLE BY INDIVIDUAL SUBJECTS ON STRATEGIES RATED BY THE R SCORE SCALE AS 1 AND 2, AND 9 AND 10

Condition	A Strategies rated low (1 and 2)	B Strategies rated high (9 and 10)
Random (0,3,16)	1.61	
	2.49	
	3.09	
	1.89	
	20.09	
Median	2.79	
(I,3,13)	1.89	
	3.67	
	4.90	
	1.57	
Median	4.285	
(I,9,13)		9.03
		10.18
		3.15
		7.56
Median		8.245
$(III,3,23)_c$	2.61	6.32
$III,9,23)_c$		5.44
		6.96
		7.31
Median		6.96
$(III,3,23)_d$	8.58	
	1.24	
	1.54	6.19
	3.22	
Median	2.38	6.19
$(III,9,23)_d$		4.05
		5.41
		3.13
		4.76
		14.96
		4.20
Median		4.48
Median—All groups	2.55	6.19

strategies which have been rated low are based upon a counting device of some sort; those rated high, either parallel or are equal to the experimenter's grammar.

Several things seem obvious from Table 9. The poorer strategies go with the smaller vocabulary, not with the larger one (see Column A). In the strategies rated high (see Column B) there is either no subject, or only one subject, per condition under a vocabulary of three words. Also note that a subject spends more than twice as much average time per word in the better strategies as a subject using the less efficient ones.

Discussion

The Number of Different Alternatives (Vocabulary). The results indicate that the size of the vocabulary *does* affect pattern conception. We notice that when nine different words were used the subject was able to decode and classify information more readily than when only three different words were used. Since pattern conception is a matter of seeing meaningful relations among the parts of a set of messages, the function of the number of different alternatives should be to convey information to the decoder about these relations by pointing up similarities and differences among the messages. Information about similarity is usually conveyed through repetition, and about the difference by the use of a different symbol. When the number of alternatives is too small, the encoder is forced to use the same symbol repeatedly even when the messages are different. His problem is initially the same as that of an artist who has to paint a multicolored landscape but who only has one, or maybe two, colors to paint with. When the number of alternatives is small, for example, three as in some of our languages, there is interference in grouping the messages because of the exaggerated similarities among the sequences. With a larger number of alternatives, for example, larger than nine, one can again expect an interference in grouping the messages, but this time the interference will be due to the exaggerated differences among the sequences. This again will tend to obscure the pattern. Perhaps between these two extremes of too few and too many alternatives there is a number which serves the ideal function of emboldening the pattern for the decoder, other things being equal.

Individual data for 3-group (any grammar group tested with three alternatives) show that when faced with haphazard similarities subjects resort to some kind of a counting device. Five subjects out of 30 in the 3-group followed the Length strategy (counting) and 12 subjects in the same group used counting of some other form. Thus a total of 17 subjects out of 30 followed some kind of a counting device. None of the subjects in the 9-group used the Length strategy and only 5 subjects used some shade of the counting strategy. Seventeen subjects in the 3-group as opposed to 5 in the 9-group is not too small a difference to overlook. Subjects in the 9-group, we suggest, can decode and recode the information better, hence their strategies are more closely related to the kind of information that is contained within the set. The strategies of these subjects, therefore, correspond more nearly to the grammar of the language.

In Grammar III we notice that although subjects spend the same amount of average time under all four conditions they do not employ equally good strategies in all the four conditions. Strategies used by the 9-group are more successful than the ones used by the 3-group. Thus once again evidence points to the number of different alternatives as a critical variable.

Now turn to Conditions B and D in Table 7 to make an intralanguage comparison within Grammar I. The results suggest that there is an interaction effect between vocabulary and the number of instances for each recursive loop in a basic string. Although the number of alternatives remains three, performance improves with an increase in the number of instances (see D). With four instances for each loop, subjects perform almost as well as subjects of (I,9,13); they take more time with

the set, however. Whether this increase in time for the (I,3,23) group is due to the larger size of the set or because it takes longer for the 3-group to find the pattern than it takes the 9-group, it is hard to say. Since groups (I,3,13) and (I,3,23) take the same amount of time for different sized languages, it could be argued that set size is less critical than the vocabulary factor.

Structural Similiarity (Grammar). The other notable parameter in pattern conception seems to be the structural similarity of sequences. In their studies on visual pattern recognition, Fitts et al. (1956) raised the problem of "different types and degrees of redundancy and of examining the effects of complexity of variables influencing pattern recognition" (p. 9). Although the measure of redundancy does not apply to our materials, as pointed out earlier, the notion of repetitiveness does. By substituting the phrase structural similarity for redundancy in the quotation from Fitts et al. we can state our own problem thus: What is the effect of the type and degree of structural similarity on pattern conception?

Similarities among sequences depend upon two kinds of repetitiveness: the repetition of words because of the limited vocabulary used and repetition due to the grammatical structure of the strings. In the previous section we discussed repetitiveness based upon the number of different alternatives. Structural repetitiveness was defined in terms of the number of basic strings contained in each language and the number and placement of the loops within the basic formulas. A particular grammar selects a particular type of repetitiveness. The problem arising from such selection is: Is the pattern of similarity and difference more difficult in one set than in another?

In scaling the languages for structural similarity, one could begin with the zero condition in grammatical rules, as in the random set of strings, where rules of formation are absent. Results show that in the absence of structural similarities among the strings a subject searches for peripheral similarities and he settles for the simpler strategies such as those of Length or of Initial similarity. Peripheral strategies, however, are not solely the function of nongrammatical structure. We noted earlier, that when the number of alternatives is small, as in (I,3,13), that even when the language has a grammar, the grammar of the set becomes so obscure through word repetition that a subject treats the set as a random one. In such an instance, the strategies used for a grammatical set will approximate those of the random set.

When we turn to the finite state systems, which of course are grammatical in structure, structural similarities consist in the ***number*** and ***kind*** of formulas contained within the grammars as explained before. Some questions that now arise are: (*a*) When a system is enlarged by adding one or more parts, do we increase variance or its invariance or both? Thus, are three formulas easier to handle than five? (*b*) Is the quality rather than an increase in the number of parts a critical factor in strategy construction? Are five rules with replication better than five rules without replication? More specifically, is Grammar (III,c) more simple than (I,3) and (III,d) or is it more complex?

The results do not give any clear indication either as to the degree or the directionality of the grammar effects. In the Exploratory Study, the data indicate some improvement in performance with the increase in the number of basic strings, from two strings to four, and to five (see Table 3), but the time also increased correspondingly. The better controlled experiments of the later studies give no clear indication of significant change in performance with an increase in the number of basic strings, the number of states, and the number of rules of transition. The relative difference in the grammar effects based on these factors is hard to assess for two reasons: first, there are interaction effects of other variables with grammar; second, there is no adequate measure of structural similarity. Since

grammatical similarity is more structural than peripheral, it can easily elude a subject but unless a subject takes advantage of grammatical repetitiveness implicit in the language, he fails to see structure. How well he will work with the grammar of the language will depend, in part, upon the way the messages have been encoded and, in part, upon how often they have been repeated. Without a sufficient number of different alternatives to indicate similarities and differences among the parts of the system, and without an adequate number of repetitions of the messages, the grammar of the system will remain obscure to a subject and its effect almost negligible upon his recoding behavior. On the other hand, when a subject does figure out the grammatical principle in any one subset, it seems to matter less how many subsets there are in the entire system. It then becomes simply a matter of transfer of a principle. Essentially, we can argue that there is only one basic principle of variance and invariance involved in a finite state system. All the formulas—either one, or three, or five, or any number—are merely variations of this single principle of additions or deletions of symbols to a basic, invariant, sequential pattern. Usually, a subject applies one method of classification fairly consistently to all instances in the set. If this method of classification happens to approximate to the experimenter's grammar, it seems to matter little whether three formulas have been built into the structure of the language or five. And if his strategy is different from the experimenter's grammar, then it matters even less how degenerate the grammar is, because the subject has missed the point of the grammar, anyhow.

We would like to venture another guess to explain the absence of a clear difference in the Grammar groups as measured by the kind of strategies subjects used. Presumably, each time one or more parts is added to a grammatical system, the information function is jointly and simultaneously affected by the increase in complexity and redundancy of the system. In other words, confusion and facilitation go hand in hand in grammatical redundancy and these effects tend to cancel each other. Thus more formulas neither make the task easier nor do they make it more difficult.

Although the quantitative factors of strings, states, and rules of transition seem to make little difference in performance in the present study, recursive loops create some problems for a subject. Two common difficulties pertaining to loops characterize the responses of most of our subjects. These difficulties we shall label as: the zero concept difficulty and the double-loop concept difficulty.

The zero concept: It was not infrequent to find a subject puzzled about the basic string of a formula. For example, if one of the formulas in a grammar was a (c) b b a, its basic string, a $(c)^0$ b b a, was often held as a separate card for a considerable period of time and was even placed in a separate category by some subjects. This observation seems more typical of the double-loop formula to be discussed below. When a subject had decoded a grammatical rule as the experimenter had structured it, he could easily see the arithmetical progression in the development of a loop move from 1 to $\widehat{n}$, even though quite deliberately some exponents in the progression had been left out. But it seemed difficult for a subject to see that the exponent of the loop could also be zero. Viewed historically, the zero concept was not easily come by either. Perhaps it is difficult for one to think of starting a series with nothing. Normal counting frequently begins with 1 and moves up to n. It is, therefore, quite possible that this way of counting interferes with the necessity to begin the series with the zero notation in our formulas.

The double-loop concept: The zero concept seemed even more difficult for subjects to handle in the double-loop formulas. This was so because the double-loop in itself was a hard concept for the subject. We have described the double-loop as a formula consisting of two variables, instead of just one, within an in-

variant sequence of symbols. It has been schematized thus: a (c) b c (b) a.

An R score of 8 or 9 signifies that subjects broke up the double-loop subset into three further sub-subsets as follows: a $(c)^{1-n}$ b c a, a b c $(b)^{1-n}$ a, and a $(c)^{1-n}$ b c $(b)^{1-n}$ a. The basic string was either held as a separate category, or it was placed with one of the first two sub-subsets. In watching subjects at work, it was fairly obvious to the experimenter that subjects were having considerable difficulty with the double-loop subset. They insisted upon applying the simple rule of one loop varying from 0 or 1 to n to this complex formula, and were confused with the similarity within the three sub-subsets. This can be very clearly seen in the thinking-aloud responses of our subject in the Illustrative Study. One could explain this difficulty on the basis of two characteristics of pattern response. The first, that the subject formulates his method early in the game and more or less stays with it; and the second, that the subject seeks to apply one pattern consistently throughout to all parts of the system.

Reiss pointed out that a subject constructs his "model" fairly early, and then begins to stabilize around it as he tests it with successive negative or positive instances. From the thinking-aloud responses of our subjects and on the basis of data from other studies on the evolution of strategies not reported here, Reiss seems to be right. Those subjects who initiated a counting strategy in the beginning showed a strong counting bias, and those who used the initial strategy stuck pretty closely to the left-to-right rule of comparing sequences. Exceptions did occur in the Random group, or when the method of presentation of instances was varied. In the Random set, one could find the subject trying all sorts of strategies, and finally settling on a perceptual one, when nothing else seemed to work. But in the sets where the subject had once perceived the structural order in a subset fairly early, or even later, he did not give it up. This does not mean, of course, that there is no further refinement of the strategy with an increase in the amount of evidence. It simply means that more instances only help to sharpen or modify the "model" already constructed.

This kind of fixation in strategy ties in with the point we are making about consistency in sorting. By and large, most subjects work with one method of classification for all the subsets they make. They do not seem to want to shift gears in their strategy from subset to subset. This characteristic way of responding is independent of the input, and it seems to give us some information about the human channel itself. It is, therefore, important to note that not only does a subject organize the input sequences into categories, but that he does this on the basis of a consistent principle.

Time

Since time did not seem to be a simple function of a single variable we have postponed its discussion; a few remarks are in order now. We note first, that the amount of time taken for the task is not a unique function of the size of the set although size may have something to do with it. Next we note that time seems to be positively correlated with strategy. And finally we see that time is affected by the size of the vocabulary, though not without exception.

Time and Size of Set. First, consider the relation between time and size of set. One is impressed by a lack of orderly variation in the amount of time corresponding to an increase in the size of the set. The group total time spent on a set of 13 instances was 49.56 seconds for the larger and 80.36 seconds for the smaller vocabulary. When the size of the set increases to 16 (larger vocabulary), time reduces to 39.16 seconds, and when the size increases to 18 instances (smaller vocabulary), time goes up to 67.16 seconds. Subjects take more time with a set of 13 instances developed with a vocabulary of three words, than they do with a set of 23 instances irrespective of vocabulary size. The small number of instances does not seem to make the subject's task

more manageable, but rather, it seems to make it more difficult, possibly by obscuring the pattern. A small set of instances gives him less evidence upon which to base his induction of the appropriate categories. The variables which interact with the size of the set apparently are far more critical in the time measure than the size (within limits) of the set itself. It is difficult to determine the true time function on the basis of adding just one more instance to the set. It seems one must know what that one more instance is, and also, what some of the other characteristics of the set are.

Time and Strategy. The longer time measures are correlated with the more successful strategies. Even at the risk of explaining the obvious, we shall state our point: complex strategies take more thinking and thinking takes time. We define thinking in this situation as a search for relations among parts of the information received, in order to recode this information into suitable units for use and for storage. The human link in any communication system is responsive not only to the immediate reception and use of information—its function is a far more complex one of storing this information for later use. As a finite system, the organism is confronted with an infinite amount of information from its environment; the quality of its efficiency is on trial on the basis of the recoding scheme the system envisages. How much and how well the old information can be reconstructed from the recoded units will determine the degree of efficiency of the thinking involved. A structural strategy makes the reconstruction of the original message far more probable than a simple counting strategy, for instance. It is, therefore, not difficult to see why it would also take more time. Even Donders (cited in Boring, 1950, pp. 147-149) was suggesting a hundred years ago that the longer reaction times were an indication of complex processes. We might speculate at this point and add that if we could force each subject to work with the information given in the set for a longer period of time (as we forced the subject described in the Illustrative Study) by presenting the information to him one instance at a time, we would expect more subjects, possibly all, to come closer to the original structure of the grammar, despite the complexity of the stimulus conditions. For example, one subject in the (I,3,13) group spent next to the longest time of any and came up with a better solution than any of the other subjects in that group. Give a man sufficient time and he will find some way to reduce the heterogeneity of his environment to lawful change.

Concluding Observations

It is clear from the data that pattern conception is a characteristic way in which people respond to a set of linguistic sequences, irrespective of the structure of the sequences. In an "open system" such as ours (unstructured in instruction and in meeting any given criteria) the subject must define the problem himself, construct some sort of working hypothesis to deal with the information he has extracted from the input, and look for evidence to see if his hypothesis holds. He then discovers gaps in the information he has acquired, he searches for new evidence to fill the gaps, and finally he accepts his plan or rejects it if the gaps remain unfilled. Usually, though a subject may show intolerance of uncertainty, he may occasionally decide to live with his problem. The strategy or plan he settles on is partially a function of some experimental variables.

It was interesting to observe that reminiscent of the Kantian categorical imperative, our subjects seemed to be bound by a categorizing imperative. And what is more, their categorizing seems to fall within a limited range. Whether there were 13 instances in a set, or 50, and whether there were 15 subjects tested or 150, it did not seem to matter; the mean number of categories ranged between five and seven (see Figure 4). By some strange coincidence, the integer that has been "persecuting" Miller (1956) for a decade, turns up once again in our data. What does one make of it?

Figures 4, 5, and 6 all relate to the number of categories subjects make. The function in Figure 4 represents data for the mean number of categories subjects made for the different sized sets. Data show that the number of categories increases somewhat with an increase in the total number of instances in the set, but that this increase is proportionately less.

The average frequency of the different types of categories subjects make is presented in Figure 5. By the type of category is meant the number of instances subjects put into one category; for example, there may be a one-instance category, a category with two instances in it, with three, and so on. The number of instances in a category and the total number of categories are interdependent factors, of course. Either a subject prefers to make a certain number of categories, or he prefers to put a certain number of instances together in one category. The more frequent categories are those with two, three, or four instances per category.

Figure 6 treats the same data in a slightly different way. It presents individual cumulative curves for different language conditions to show the probability that a subject will make at least x categories. Note that for all conditions 50% of the subjects average between five and six categories.

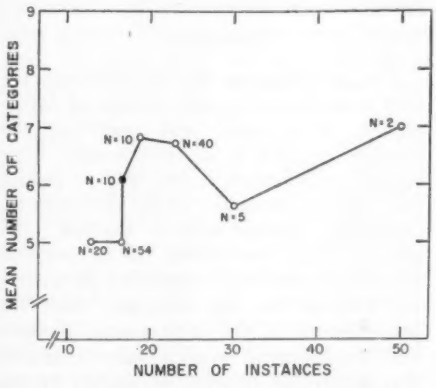


FIG. 4. Mean number of categories per size of set. (N denotes the number of subjects in each group. The single filled circle represents the Random group.)

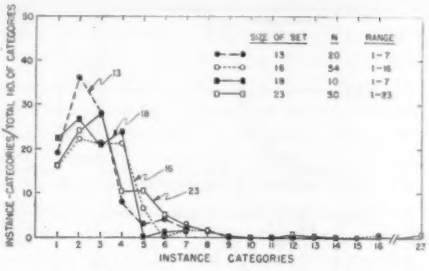


FIG. 5. Mean number of instance categories per size of set. (Instance category refers to the number of instances placed in a category. The y axis represents what percentage an instance category is of the total number of categories made in each language. For example, in the 23-instance languages, 50 subjects made a total of 338 categories, of which 54 categories were one-instance categories.)

We have stated above that although the number of instances in a set has something to do with the number of categories subjects make, it is not a determining factor. The relation of the number of basic strings in a set to the number of categories seems even less important, except when a subject sees the grammatical pattern accurately. Quite frequently subjects wished to know how many categories they were expected to make. When they were informed that the choice of number rested with them, they seemed to guide their sorting by a desire not to have "too many" categories. This was especially true of subjects employing some type of formula strategies. "It's a shame to have a one-card category" was not unusual for a subject to say when an instance remained unclassified. If, however, the choice before the subject was between consistency in sorting and the number of categories, he preferred consistency to number.

Seven, then, is still a magic number. The "number seven, plus or minus two" denotes according to Miller (1956), "the span of absolute judgment" and the "span of memory" (p. 90). We interpret the spans to mean that there is a finite number of isolated events, or of groups of events, within the neighborhood of seven, which the human link in a communication system

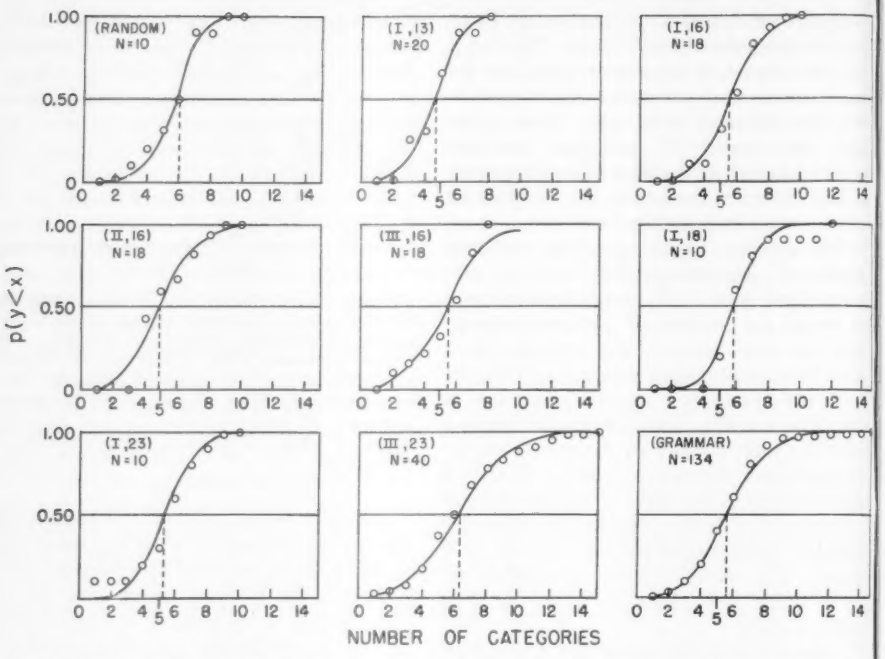


FIG. 6. The probability that a subject will make at least x categories is represented by $p(y < x)$ for nine groups of subjects. (The S-shaped functions represent theoretical curves fitted to the data on probability paper for each group separately. Circles represent actual data. Data from all Grammar groups have been combined in the last figure where $N = 134$.)

can best handle. It seems, then, that in dealing with a bulk of information the human channel is constrained by its "channel capacity," to use Miller's terminology, to process the given information within the given range of categories. It seems as though, because of the limits imposed by the memory span upon pattern conception, one might begin to develop a notion of conceptual span as well. Not only is memory limited by the number of psychological units it can handle, but pattern conception must function under the same limitation.

Although the span of immediate memory is limited by the number of items, the amount of information the items contain can be increased by a process of recoding. Pattern conception is a description of the recoding process itself. The linguistic recoding involved in the type of studies presented above is basic to the understanding of memory and thinking. The study is in its infancy; much more research is warranted.

SUMMARY

It is advantageous for an organism, for both communication and storage of information, to organize and symbolize the stimulus energies in its environment. Perceptual organization is one way to reduce the apparent variability of the environment, conceptual symbolization is another. Together, these two forms of information processing provide an organism with powerful weapons to deal efficiently with the heterogeneity of the environment. Numerous psychological studies have indicated that perceptual recoding is limited by the amount of information given; is conceptual recoding similarly related in any specifiable way to the given information?

Conceptual recoding is here discussed in terms of the discovery and formulation of rules to describe cognitive structures. Language provides one of the best illustrations; verbal recoding of experience seems to be the essence of human thinking. It is to language that one can turn, therefore, to find the materials needed for experimental study of the processes of conceptual recoding. Natural languages, however, are tremendously complicated; some of the same processes can be studied more simply with artificial language systems. In the present study finite state systems were used to simulate natural languages. They are simpler, their formal structure can be stated precisely, yet they provide the kind of sequential patterning that is characteristic of all verbal communication. Subjects were given a set of linguistic strings, random or constrained, generated by some finite state systems; they were asked to group the strings in some appropriate way. In order to perform such a task all subjects employ some intelligible, recognizable plan for sorting the strings. The number of different strategies that are observed to occur is small and subjects are easily classified according to the strategy they adopt. Some strategies are more complicated than others; they range from simple, ready-made rules for counting and alphabetizing to ad hoc inventions of rules that a subject considers implicit in the information he has been given. The former are more perceptual, the latter more conceptual in spirit, but both are ways of describing the structural rules of the system.

Pattern conception depends upon the kind and amount of the evidence supplied by the experimenter. In the present studies, the length of the messages, the number of instances, the number of alternatives, and structural similarity are some of the experimental variables by which messages were varied. The results show that the variety of symbols (number of different alternatives) with which the messages are constructed is far more critical in controlling pattern conception than is the length or the structure of the messages. Two sets of messages, identical in length, in rules of formation, and in the instruction given subjects, but differing in the number of symbols, elicit qualitatively different responses. A large variety of symbols tends to elicit the invented strategies whereas the small variety generally leads to ready-made rules. This result indicates that the subject is able to recode the given information more effectively when there is a variety in the symbols from which it is composed.

A parallel advantage is observed by increasing the amount of evidence provided to a subject for a given rule. The number of instances per rule was varied from two to three to four instances, while other factors of vocabulary, grammar, and instructions were held constant. The strategies that subjects used under each of these conditions were evaluated in terms of two statistical techniques developed specifically for the purpose of strategy analysis. As the amount of evidence increased, the strategies improved. A larger quantity of evidence facilitates recoding of the sequences in terms of variant and invariant symbols.

Earlier experiments on perception have indicated that redundancy may be of several kinds, and that the influence of redundancy upon responses varies with its quality. Although we have not applied an information measure of redundancy to our materials (the finite state systems that we used are too simple to require such statistical treatment) we have worked with a related conception of structural similarity (rules of formation) in linguistic sequences. Languages varying in number and complexity of rules of formation, but similar in length and size of vocabulary, do not yield markedly different results. The fewer rules embedded in a language do not necessarily make that language simpler. Data indicate that structural redundancy is less critical in pattern conception than is the variety in symbols or the amount of evidence. More experiments with better controls on redundancy are needed to determine the effect of this variable on pattern conception.

Two specific difficulties were found to characterize the responses of many subjects in almost all of the experimental groups. These difficulties were the concept of zero and the concept of double variation in a linguistic sequence of variant and invariant symbols. Subjects could delete or add the variable from 1 to *n* fairly easily, but they struggled with the string when the exponent of the variable was equal to zero. In the case of more than one variable in a string subjects tended to group its instances into as many subsets as there were variables in the string. This multiple-loop difficulty is not unrelated to the zero difficulty, however.

Pattern conception, apparently, is not a simple function of any one experimental variable. This point was most clearly evident from the time measures, which did not correlate consistently with any one experimental condition. It was reasonably clear, however, that the longer time measures correlated positively with the better strategies. Presumably, the central processes of recoding and thinking take time, which causes the delay in response. There was little evidence for an increase in time with the increasing length of the set.

Finally, all studies led us to conclude that, irrespective of the type of finite state system used—its structure, vocabulary, length, or anything else—the average number of categories that subjects make is five or six. Although we were not looking for this point, this result is in keeping with the findings from some studies on "channel capacity," and it indicates that thinking like memory is centrally affected by the channel capacity of the organism.

REFERENCES

ABORN, M., & RUBENSTEIN, H. Information theory and immediate recall. *J. exp. Psychol.*, 1952, **44**, 260–266.

ADAMS, O. S., FITTS, F. M., RAPPAPORT, M., & WEINSTEIN, M. Relations among some measures of pattern discriminability. *J. exp. Psychol.*, 1954, **48**, 81–84.

ANDERSON, NANCY S., & LEONARD, J. A. The recognition, naming, and reconstruction of visual figures as a function of contour redundancy. *J. exp. Psychol.*, 1958, **56**, 262–270.

ATTNEAVE, F. Some informational aspects of visual perception. *Psychol. Rev.*, 1954, **61**, 183–194.

ATTNEAVE, F. Symmetry, information and memory for patterns. *Amer. J. Psychol.*, 1955, **68**, 209–222.

ATTNEAVE, F. *Applications of information theory to psychology: A summary of basic concepts, methods, and results.* New York: Holt, 1959.

BARTLETT, F. *Thinking: An experimental and social study.* New York: Basic Books, 1958.

BORING, E. G. *A history of experimental psychology.* New York: Appleton-Century-Crofts, 1950.

BROGDEN, W. J., & SCHMIDT, R. E. Acquisition of a 24-unit verbal maze as a function of the number of alternative choices per unit. *J. exp. Psychol.*, 1954, **48**, 335–338. (a)

BROGDEN, W. J., & SCHMIDT, R. E. The effect of number of choices per unit of a verbal maze on learning and serial position errors. *J. exp. Psychol.*, 1954, **47**, 235–240. (b)

BRUNER, J. S. GOODNOW, J. J., & AUSTIN, G. A. *A study of thinking.* New York: Wiley, 1956.

CARMICHAEL, L., HOGAN, H. P., & WALTER, A. A. An experimental study of the effect of language on the reproduction of visually perceived forms. *J. exp. psychol.*, 1932, **15**, 73–86.

CHOMSKY, N. *Syntactic structures.* Hague: Mouton, 1957.

CHOMSKY, N., & MILLER, G. A. Finite state languages. *Inform. Control*, 1958, **1**, 91–112.

FITTS, P. M., WEINSTEIN, M., RAPPAPORT, M., ANDERSON, NANCY S., & LEONARD, J. A. Stimulus correlates of visual pattern recognition: A probability approach. *J. exp. Psychol.*, 1956, **51**, 1–11.

FRENCH, R. S. The discrimination of dot patterns as a function of number and average separation of dots. *J. exp. Psychol.*, 1953, **46**, 1–9.

FRENCH, R. S. The effect of instruction on the length-difficulty relationship for a task involving sequential dependency. *J. exp. Psychol.*, 1954, **48**, 89–97. (a)

FRENCH, R. S. Identification of dot patterns from memory as a function of complexity. *J. exp. Psychol.*, 1954, **47**, 22–26. (b)

FRENCH, R. S. Pattern recognition in the presence of visual noise. *J. exp. Psychol.*, 1954, **47**, 27–31. (c)

GOLDSTEIN, K., & SCHEERER, M. Abstract and concrete behavior: An experimental study with special tests. *Psychol. Monogr.*, 1941, **53**(2, Whole No. 239).

HANFMANN, EUGENIA, & KASANIN, J. A method for study of concept formation. *J. Psychol.*, 1937, **3**, 521–540.

HEIDBREDER, EDNA. The attainment of concepts: II. The problem. *J. gen. Psychol.*, 1946, **35**, 191–223.

HOVLAND, C. I. The generalization of conditioned responses: I. The sensory generalization of conditioned responses with varying frequencies of tone. *J. gen. Psychol.*, 1937, **17**, 125–148. (a)

HOVLAND, C. I. The generalization of conditioned responses: II. The sensory generalization of conditioned responses with varying intensities of tone. *J. genet. Psychol.*, 1937, **51**, 279–291. (b)

HOVLAND, C. I. A "communication analysis" of concept learning. *Psychol. Rev.*, 1952, **59**, 461–472.

HULL, C. L. Quantitative aspects of the evolution of concepts. *Psychol. Monogr.*, 1920, **28**(1, Whole No. 123).

JOHNSON, D. M. *The psychology of thought and judgment.* New York: Harper, 1955.

KOFFKA, K. *Principles of gestalt psychology.* New York: Harcourt, Brace, 1935.

KÖHLER, W. *Gestalt psychology.* New York: Liveright, 1929.

KÖHLER, W. *The place of value in a world of facts.* New York: Liveright, 1938.

KÖHLER, W. *Gestalt psychology: An introduction to new concepts in modern psychology.* New York: Liveright, 1947.

KRECHEVSKY, I. "Hypotheses" vs. "chance" in the pre-solution period in sensory discrimination-learning. *U. Calif. Publ. Psychol.*, 1932, **6**, 27–44.

LUCE, D. R. A survey of the theory of selective information and some of its behavioral applications. New York: Bureau of Applied Social Research, 1954. (Behavioral Models Project (NR 042-115) CU-23-54-NONR-266621-BASR).

MILLER, G. A. *Language and communication.* New York: McGraw-Hill, 1951.

MILLER, G. A. What is information measurement? *Amer. Psychologist*, 1953, **8**, 3–11.

MILLER, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychol. Rev.*, 1956, **63**, 81–97.

MILLER, G. A. Free recall of redundant strings of letters. *J. exp. Psychol.*, 1958, **56**, 485–491.

MILLER, G. A., & CHOMSKY, N. Pattern conception. Paper delivered at Conference on Pattern Detection at University of Michigan, 1958.

MILLER, G. A., & SELFRIDGE, JENNIFER A. Verbal context and the recall of meaningful material. *Amer. J. Psychol.*, 1950, **63**, 175–185.

MOSTELLER, F., & BUSH, R. R. Selected quantitative techniques. In G. Lindzey (Ed.), *Handbook of social psychology*. Vol. I. New York: Addison-Wesley, 1954. Ch. 8.

OSGOOD, C. E. *Methods and theory in experimental psychology*. New York: Oxford Univer. Press, 1956.

QUASTLER, H. (Ed.) *Information theory in psychology: Problems and methods*. Glencoe: Free Press, 1955.

RAZRAN, G. H. S. Studies in configurational conditioning: V. Generalization and transposition. *J. genet. Psychol.*, 1940, **56**, 3–11.

RAZRAN, G. H. S. Stimulus generalization of conditioned responses. *Psychol. Bull.*, 1949, **46**, 337–365

REISS, D. Subjective models of a simple finite-state grammar: A study of their attainment and characteristics. Unpublished honors thesis, Harvard University, 1958.

RUBENSTEIN, H., & ABORN, M. Immediate recall as a function of degree of organization and length of study period. *J. exp. Psychol.*, 1954, **48**, 146–152.

SHANNON, C. E., & WEAVER, L. W. *The mathematical theory of communication*. Urbana: Univer. Illinois Press, 1949.

SMOKE, K. L. An objective study of concept formation. *Psychol. Monogr.*, 1932, **42** (4, Whole No. 191).

SPENCE, K. W. The solution of multiple choice problems by chimpanzees. *Comp. psychol. Monogr.*, 1939, **15**, No. 3.

THORNDIKE, E. L. *Human learning*. New York: Century, 1931.

VON EHRENFELS, C. On gestalt qualities. *Psychol. Rev.*, 1937, **44**, 521-524.

WERTHEIMER, M. *Productive thinking*. New York: Harper, 1945.

WITKIN, H. A. I. "Hypotheses" in rats: An experimental critique. *J. comp. Psychol.*, 1940, **30**, 457–482.

WOODWORTH, R. S. *Contemporary schools of psychology*. New York: Ronald, 1948.

(Received April 11, 1960)

APPENDIX A

Agreement-Disagreement Ratio

The Agreement-Disagreement ratio is a statistic devised to measure the amount of correlation between sortings made by any two persons. The technique of calculating the amount of agreement between different sortings is as follows:

Consider the sortings made by any two subjects X and Y:

X—(1 2 3 4 5) (6 12 16) (9 14)
(7 8 10 11 13 15)

Y—(1 2 4 5) (3) (6 10 12 15)
(7 8 9 11 13 14) (16)

Step 1. Construct sets from the intersection of these two classifications:

Intersection (1 2 4 5) (6 12) (7 8 11 13)
(9 14) (10 15) (3) (16)

The intersection is constructed in such a way that every set in the data for X and Y can be reconstructed by combining sets from the intersection, and yet the sets in the intersection are as large as possible. Otherwise said, every distinction among the strings that either X or Y indicated is preserved in the intersection.

Step 2. Calculate the number of agreements for each set. For example, if a set contains four strings there are $3 + 2 + 1 = 6$ pairs of strings that agree with one another in the sense that they are placed in the same class and considered to be equivalent. Below are tabulated the number of pairs of strings that agree as a function of the size of set.

size	*number agreements*
1	0
2	1
3	3
4	6
5	10
x	$x(x-1)/2$

Thus, for example, a set of 11 strings would yield 55 pairs that were in agreement. If size is x then the number of agreements is given by the equation $A = x(x-1)/2$. So the second step is to add up the number of agreements in this sense for X, Y, and the intersection:

subject X		*subject Y*		*intersection*	
5	10	4	6	4	6
3	3	1	0	1	0
2	1	4	6	2	1
6	15	6	15	4	6
	29	1	0	2	1
			27	2	1
				1	0
					15

Step 3. Calculate the A-D ratio from these three numbers, 29, 27, 15, by substituting them into the following equation:

$$r = 1 - \frac{x + y - 2\,\text{Intersection}}{60}$$

$$r = 1 - \frac{29 + 27 - 2\,(15)}{60} = 1 - \frac{26}{60}$$

$$= 1 - .443 = .567$$

APPENDIX B

Rating Scale for Subjects' Strategies

The second column gives the average scores for the ratings assigned to 20 strategies by five judges. The first column rounds off the average scores for purposes of establishing a 10-point rating scale.

Final Rating Score	Average Score[a]	Kind of Strategy
1	1.0	General Size
	1.5	Length
2	2.4	$Initial_1$, or $Middle_1$, or $Final_1$
	2.5	Alphabet
3	2.7	One striking word with Initial or Final one
	3.1	$Initial_{(2 \text{ or } n)}$, $Middle_{(2 \text{ or } n)}$, $Final_{(2 \text{ or } n)}$
	3.1	$Initial_1$-$Final_1$
	3.4	Length plus Alphabet
	3.4	$Initial_1$ plus Alphabet or Initial plus Length
	3.6	Rhythm
4	3.8	Mixed methods for entire set
	3.8	Mixed methods by groups
	4.3	Alternating word or complex counting
5	5.3	Symmetry and Rhythm
	5.6	Initial plus complex substitution or ordering
6	6.2	$Initial_{(2 \text{ or } 3)}$-$Final_{(2 \text{ or } 3)}$
7	7.4	Maximum Identity
8	8.2	Initial-Middle-Final
9	9.0	Less than basic string (breaking up the formula by loops)
10	10.0	Formula (same as experimenter's grammar)

[a] The steps between scores are not necessarily equal.